Tampere University

Eero Halonen

CALIBRATION METHODS OF CHARACTERIZATION LENS FOR HEAD MOUNTED DISPLAYS

Faculty of Engineering and Natural Sciences
Master of Science Thesis
November 2019

ABSTRACT

Eero Halonen: Calibration Methods of Characterization Lens for Head Mounted Displays
Master of Science Thesis
Tampere University
Science and Engineering
November 2019

This thesis concerns the calibration, characterization and utilization of the HMD Eye, OptoFidelity's eye-mimicking optical camera system designed for the HMD IQ, a complete test station for near eye displays which are implemented in virtual and augmented reality systems. Its optical architecture provides a 120 degree field of view with high imaging performance and linear radial distortion, ideal for analysis of all possible object fields. HMD Eye has an external, mechanical entrance pupil that is of the same size as the human entrance pupil.

Spatial frequency response (the modulation transfer function) has been used to develop sensor focus calibration methods and automation system plans. Geometrical distortion and its relation to the angular mapping function and imaging quality of the system are also considered.

The nature of the user interface for human eyes, called the eyebox, and the optical properties of head mounted displays are reviewed. Head mounted displays consist usually of two near eye displays amongst other components, such as position tracking units. The HMD Eye enables looking inside the device from the eyebox and collecting optical signals (i.e. the virtual image) from the complete field of view of the device under test with a single image.

The HMD Eye under inspection in this thesis is one of the 'zero' batch, i.e. a test unit. The outcome of the calibration was that the HMD Eye unit in this thesis is focused to 1.6 m with an approximate error margin of $\pm$10 cm. The drop of contrast reaches 50% approximately at angular frequency of 11 cycles/degree which is about 40% of the simulated values, prompting improvements in the mechanical design. Geometrical distortion results show that radial distortion is very linear (maximum error of 1%) and that tangential distortion has a diminishable effect (0.04 degrees of azimuth deviation at most) within the measurement region.

Keywords: Calibration, modulation transfer function, geometric distortion, human eye, head mounted display

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

TIIVISTELMÄ

Eero Halonen: Päähinenäyttöjä karakterisoivan linssisysteemin kalibrointimetodit
Diplomityö
Tampereen yliopisto
Teknis-luonnontieteellinen
Marraskuu 2019

Tämän diplomityön aiheena on OptoFidelityn päähinenäytöille suunnatun ihmissilmää jälijittelevän HMD Eyen kalibrointi, karakterisointi ja hyödyntäminen. Laitetta käytetään osana HMD IQ -testausasemaa. Kyseinen optinen arkkitehtuuri mahdollistaa 120 asteen katselukulman laadukkaalla piirtokyvyllä ja lineaarisella radiaalisella vääristymällä, joka on ihanteellinen ominaisuus havaittavien kohdekenttien tarkastelua varten. HMD Eyessä on irrallinen, ihmissilmän pupillin kokoa vastaava, mekaaninen sisäänmenopupilli.

Avaruustaajuusvasteen (modulaationsiirtofunktion) avulla on kehitetty kameran fokusointiin tarvittavia menetelmiä ja sen automatisoinnin mahdollistavan järjestelmän suunnitelmia. Geometrisen vääristymän yhtymäkohdat kulma-avaruuden kuvaamisfunktioon ja piirtämisen suorituskykyyn otetaan esille.

Päähinenäyttöjen fyysinen rajapinta ihmissilmien käyttöliittymälle, eyebox, ja muut optiset ominaisuudet ovat myös käsittelyn kohteena. Päähinenäytöissä on tavallisesti omat monitorit kummallekin silmälle, sekä muita komponentteja kuten asennonseurantayksikköjä. HMD Eye mahdollistaa päähinenäyttöjen lähettämän optisen signaalin (virtuaalikuvan) keräämisen täydestä katselukulmasta yhteen digitaaliseen kuvaan eyeboxin välityksellä.

Diplomityössä käytetään "nollasarjan"HMD Eye -yksikköä, toisin sanoen testikappaletta. Kalibroinnin lopputuloksena saatiin 1,6 m etäisyyteen fokusoitu HMD Eye arviolta $\pm$10 cm:n virhemarginaalilla. Kontrasti putoaa 50 %:iin kulma-avaruuden taajuudella 11 jaksoa/aste, joka on noin 40% simuloidusta lukuarvosta, mikä kehottaa mekaaniikan parannustoimenpiteisiin. Geometrisesta vääristymästä saadut tulokset osoittavat, että radiaalinen vääristymä on hyvin lineaarinen (suurin virhe 1 %). Tangentiaalisella vääristymällä on mitätön vaikutus (korkeintaan 0,04 asteen poikkeama atsimuuttikulmassa) mittausalueella.

Avainsanat: Kalibrointi, modulaationsiirtofunktio, geometrinen vääristymä, ihmissilmä, päähinenäyttö, virtuaalitodellisuus

PREFACE

This Master of Science thesis has been written for OptoFidelity. My keen interest in photonics, optics and related applications in technology has been supported by OptoFidelity. I am grateful for the opportunities and praxis this project has provided me.

The HMD IQ project evoked my curiosity with its various interesting details. The HMD Eye is definitely the most significant territory for me and has influenced and helped me advance in my field of expertise. I want to thank Sakari Kapanen and Ville Hautala for developing the software for the core calibrations processes and Sounak Bhattacharya and Tuomo Ojaluoma for the help with processing the headset under investigation. I also want to thank Ilkka Alasaarela, Kimmo Jokinen and Kalle Ryynänen for their guidance. I'm very much obliged to the whole HMD IQ team for their help.

I would like to express my deepest gratitude to my mentor figure and thesis supervisor Janne Simonen, who has dedicated his time for listening to my deliberation during my studies, first years at OptoFidelity and of course this thesis. I also want to thank Juha Toivonen for reviewing this thesis and sharing his interest and ideas.

This thesis would not have been possible without the support and nurturing of my family, friends and fiancée Pinja. Your warm support and selfless encouragement is something I have always been able to rely on.

Tampere, 3rd November 2019

Eero Halonen

CONTENTS

1 Introduction 1
2 Optical Principles of Imaging Systems 3
2.1 Wave Mechanics 3
2.2 Ray Optics 4
2.3 Lens Systems 6
2.4 Stops and Pupils 8
2.5 Diffraction 9
3 Imaging System Parameters 11
3.1 Field of View and Resolution 11
3.2 Chromatic Aberration 11
3.3 Contrast and Modulation Transfer Function 12
3.4 Geometric Distortion 16
3.5 Flatfield and Dark current 21
3.6 Optical Axis and Rotation Symmetry 24
4 Near Eye Display Parameters and Quality Assurance 25
4.1 Human Eye 25
4.2 Display Technologies 27
4.3 Imaging Parameters in Near Eye Displays 28
4.4 Eyebox 28
4.5 Vergence and Eye Tracking 30
4.6 Defects and Factory Environments 31
5 HMD Eye Lens and Camera 33
5.1 Design Principles 33
5.2 Competitor Products 35
6 HMD Eye Calibration Methods 36
6.1 Focus and Modulation Transfer Function 36
6.2 Measuring and Correcting Geometrical Distortion 41
6.3 Dark Frame and Flatfield Correction 42
6.4 Exposure and Power Linearity 43
6.5 Optical Axis and Rotation Symmetry 44
6.6 Calibration Automation 44
7 Experimental Results 48
7.1 Results for HMD Eye 48
7.2 Results for Virtual Reality Headset 53
8 Analysis of Experimental Results 57

8.1 Analysis of HMD Eye Measurement Results 57
8.2 Analysis of Virtual Reality Headset Measurement Results 66
9 Conclusion 72
References 74

LIST OF SYMBOLS AND ABBREVIATIONS

C	Optical center
C_i	Corrected pixel value
D	Darkframe pixel value
F	Flatfield pixel value
F_max	Maximum flatfield pixel value
F_o, F_i	Front and back focal points
H_1, H_2	First and second primal point
I	Image point
I_max, I_min	Maximum and minimum intensity
J_1	Bessel function of the first order
K_ca	Relative lateral chromatic aberration
K_dist	Relative geometric distortion
K_m	Luminance scaling factor
L_e	Radiance
L_v	Luminance
M_L	Longitudinal magnification
M_T	Transverse magnification
N_1, N_2	First and second nodal point
O	Object point
O_i	Original pixel value
P_i	Chief ray intersection points
R, T	Radially and tangentially distorted image points
R_min	Minimum from source or destination distance respect to aperture
V_1, V_2	Lens vertices
κ	Spatial period
λ	Wavelength
$\mathbf{k}$	Spatial frequency
$\mathbf{k}_\mathsf{Nyquist}$	Nyquist frequency
$\mathbf{v}$	Speed

ν	Frequency
θ	Angle in respect to surface normal
a	Greatest width of the aperture
c	Speed of light in vacuum
d_{e}	Eye relief
d_{i}	Interpupillary distance
f	Focal length
f_{eff}	Effective focal length
i	Imaginary unit
k	Wavenumber
k_i	Radial distortion coefficients
n	Refractive index
p_i	Tangential distortion coefficients
q_{1}	Airy disk radius
r	Tangent of object angle
s_{o}, s_{i}	Object and image distances in respect to the primal planes
s_i	Thin prism distortion coefficients
t	Time
u, v	Horizontal and vertical image coordinates
w	Slope of a line
x, y	Horizontal and vertical object coordinates
x', y'	Horizontal, vertical object angular coordinates
x'', y''	Horizontal and vertical distorted object angular coordinates
x_{o}, x_{i}	Object and image distances in respect to the focal planes
y_{o}, y_{i}	Object and image transverse heights
1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
AR	Augmented Reality
CIE	Commission Internationale de l'Eclairage
cpd	Cycles Per Degree
D65	CIE Standard Illuminant D65
DFT	Discrete Fourier Transform
ESF	Edge Spread Function

FIR	Finite Impulse Response
FWHM	Full Width Half Maximum
GRIN	Graded Refractive Index
HMD	Head Mounted Display
ISO	International Organization for Standardization
ISP	Integrating Sphere Photometer
LCD	Liquid Crystal Display
LCOS	Liquid Crystal on Silicon
LSF	Line Spread Function
MR	Mixed Reality
MTF	Modulation Transfer Function
MTF20	Spatial frequency where modulation depth (contrast) has decreased to 20%
MTF50	Spatial frequency where modulation depth (contrast) has decreased to 50%
NED	Near Eye Display
OECF	Opto-electronic Conversion Function
OLED	Organic Light Emitting Diode
OTF	Optical Transfer Function
ppd	Pixels Per Degree
PTF	Phase Transfer Function
RGB	Red Green Blue color model
ROI	Region of Interest
SNR	Signal-to-Noise Ratio
SOP	Standard Operation Procedure
TAU	Tampere University
VAC	Vergence-Accommodation Conflict
VR	Virtual Reality

1 INTRODUCTION

There is a ever-increasing demand for products that reshape, replace and augment reality. Over 5.5 million virtual reality (VR) device units were shipped in 2018 and the number for augmented reality (AR) devices is expected to grow from 0.25 million to over 0.5 million in 2019 [1].

The human sensory system can detect many forms of signals. VR and AR devices mainly consider visual senses (amongst sound), but other senses can be stimulated as well.

VR devices replace the visual reality humans observe while AR devices add information to the present reality. Since human beings are very sensitive to these signals, immersion is very difficult to achieve. Therefore, it is crucial to test systems that try to achieve this goal.

Measuring and testing these systems in a universal manner can be achieved with optical metrology equipment that behaves similar to the human eye. An arbitrary fish-eye lens is not suitable. The reasoning behind this is that the size of the human entrance pupil is very small, approximately 3.6 mm. Evidently, very special optical architecture is demanded. Since the output of a system is dependent on the sum of its subsystems, the metrology equipment has to be well known, i.e. calibrated, in order to address the system under inspection.

In this thesis, the calibration, characterization and utilization of the HMD Eye, OptoFidelity's eye-mimicking optical camera system, are covered. HMD Eye is a part of the HMD IQ, a complete test station for near eye displays, which are implemented in head mounted displays such as virtual and augmented reality systems.

In terms of this project, physical systems are predominantly modeled and analyzed in the spatial domain. Temporal phenomena are mostly absent, both in the case of HMD Eye or head mounted displays. For temporal domain performance analyses of head mounted displays, OptoFidelity provides the Buddy system.

HMD Eye provides a 120 degree field of view with high imaging performance and linear radial distortion which is ideal for analysis of all possible object fields. HMD Eye has an external, mechanical entrance pupil that is of the same size as the human entrance pupil. This enables characterization of head mounted displays that have an interface designed for human pupil size.

In order to address the more specific topics related to the subject, basic theory of optics

is covered. The main point of view are the electromagnetic wavelets, i.e. light rays. The basic properties of optical systems that can be derived with light rays are described.

After this, more specific optical phenomena are reviewed. These are connected to the input (objects) and output (images) of optical systems in general. This gives tools to articulate important principles related to the topic and upgrades of the project.

Finally, measurement and analysis results of HMD Eye calibration and characterization of a device under test (DUT) are presented. The goal is to review the test unit HMD Eye and the calibration method plans. Characterization of the given DUT serves as a reference how well these plans are implemented in practice.

2 OPTICAL PRINCIPLES OF IMAGING SYSTEMS

At first, the fundamental physical properties of light are considered. The primary result is the concept of light rays, which are electromagnetic wavelets that together constitute the output of a given light source.

2.1 Wave Mechanics

According to wave-particle duality [2], light can be quantified as particles and as electromagnetic wave motion oscillation. Wave mechanics is exerted in this thesis since ray optics, i.e. basis of imaging, is heavily based on it. When light is considered as wave motion, the basic rule of wave mechanics applies [2]:

$$\lambda\nu = \mathbf{v}, \tag{2.1}$$

where λ is the wavelength (the physical separation of two consecutive maxima at any given time), ν is the frequency (inverse of the time how long it takes for the maximum to reach the location of the next maximum at $t = 0$) and $\mathbf{v}$ is the speed of the wave. The speed of light in vacuum is denoted as c and it is a universal physical constant [2]. When light proceeds through a medium the speed of light depends on the material and the environment. This result is encapsulated in the form of refractive index of light [2]:

$$n = \frac{c}{\mathbf{v}}. \tag{2.2}$$

Refractive index is not only a function of material and the environment, but also the frequency of the light itself [3]. This becomes evident when considering the phenomena on a quantum level: as energy levels of the material are quantified, surely different prospects of events are bound to arise. Difficulties are met because of this phenomenon in many applications, as to be discussed in Chapter 4. Wavelength is often used instead of frequency in describing this dependence, but herein lies a risk of mixing up the wavelength in material and in vacuum.

When plane wave light reaches a surface with an angle θ_1 in respect to the surface normal, each point of the wavefront passes through the surface at different times. However, the frequency of the light cannot change (energy is a constant in this case) and the wave-

fronts have to be that of a plane wave (since each point in the wavefront is treated equally, but at a different, linearly separated time). This means that the angle (now θ_2) and the wavelength of the wavefront have to change (in equation 2.1 ν is a constant and varying $\mathbf{v}$ affects λ) so that each of the points in the same phase (in a certain period) remain in the same plane. With this argument Snell's law 2.3 can be derived [2]:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2, \tag{2.3}$$

where n_1 is the refractive index of the medium before the surface and n_2 that of after it. The phenomenon described in Equation 2.3 is called refraction.

When a planar wave interacts with nonplanar surfaces, the output is not a plane wave. If the plane wave is conceptually split into small enough parts, these can be considered to interact with an approximately planar surface. This is utilized when deriving equations related to lens optics.

2.2 Ray Optics

Ordinary objects in real life emit light in all directions. This makes analysis of optics more extensive, particularly before scientific computation. As stated in section 2.1, the overall wave should be split into small, approximately planar wavelets. These wavelets are called light rays (just rays from now on).

When rays are interacting with material surfaces, Snell's law applies. When a medium (with $n > 1$) is confined within two spherical surfaces, and the thickness d of the material can be approximated as zero, the system is called a thin lens. The line that intersects with both centers of curvature (of spherical surfaces) is called the optical axis. When viewing the optical axis from an orthogonal point of view, the left side of the lens in Figure 2.1 is called the object space where light sources, also known as objects, are placed. A point that is a source of rays that reach the lens, is called an object point.

Rays that originate from the same object point are incident to the lens interface at different angles (unless the object is infinitely far from the lens) and intersect it at different points. Snell's law is applied to each of the rays at both surfaces. If we make an approximation that all the ray angles in respect to the optical axis are very small (i.e. the sine of the angle is the angle itself and cosine equals one), the result is that the rays converge to a single point, called the image point, on the right side of the lens, the image space. This is called the paraxial approximation (also known as Gaussian optics or first order approximation based on the Taylor series expansion, which is a representation of functions as an infinite sum of terms) and it relates the object and image distances s_o and s_i (the perpendicular distance of object and image points with respect to the lens) in a closed form. If there is point-to-point correspondence in the optical system, it is said to be stigmatic. [4]

This result is applicable to object points that are off-axis, meaning that the object is lo-

cated at a distance from the optical axis named the object height y_o. Similarly, the image point distance from the axis is named as the image height y_i. These are portrayed in Figure 2.1.

Incoming rays from object point O that are parallel to the optical axis refract so that they intersect the axis at back focal point F_i. Rays that intersect the front focal point F_o leave the lens parallel to the optical axis. With a thin lens, the perpendicular distance of F_i and F_o to the lens is the same, named as the focal length f of the lens. Rays that intersect the optical center of the lens (physical center in the case of a thin lens) retain their angle. As stated earlier, these three special rays (and all rays that reach the lens) converge to the image point I. [4]

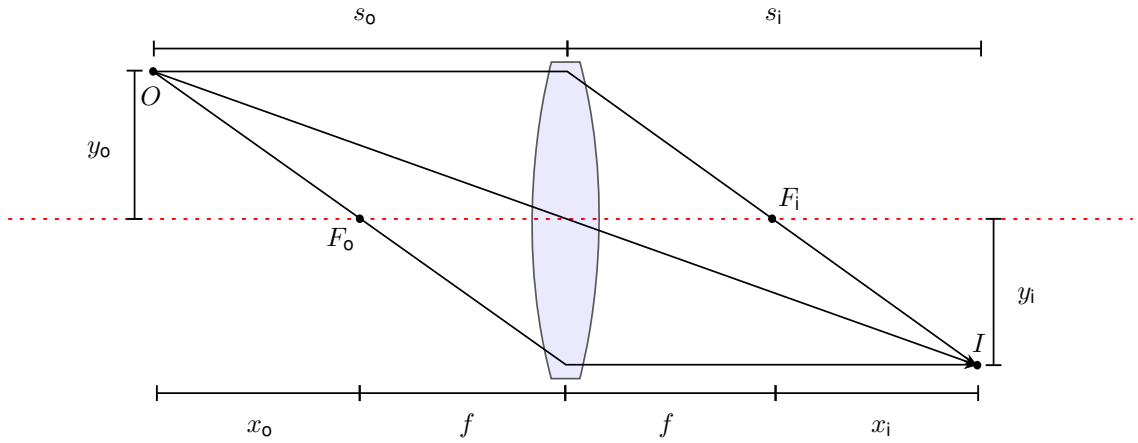


Figure 2.1. *A single thin lens imaging an object point into an image point.*

The object and image distances can also be defined respect to the focal point (F_o and F_i respectively), notated as x_o and x_i as seen in Figure 2.1. From Figure 2.1 many relations can be derived, put together in equation 2.4:

$$\frac{y_o}{|y_i|} = \frac{f}{s_i - f} = \frac{s_o}{s_i} = \frac{x_o}{f} = \frac{f}{x_i}. \tag{2.4}$$

Arguably the most used equation derived from Relation 2.4 is the Imaging equation:

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}, \tag{2.5}$$

where the inverse of the focal length is also known as the refractive power.

Depending on the object distance and the focal length, the image distance could also be negative. This means that the rays diverge in the image space, but their extensions in the object space converge to a point. This is called a virtual image. Another form of the

imaging equation is the Newtonian lens formula:

$$x_{\mathrm{o}} x_{\mathrm{i}} = f^2. \tag{2.6}$$

The relation between the image and object height is named the transverse magnification M_T [4]:

$$M_{\mathrm{T}} = \frac{y_{\mathrm{i}}}{y_{\mathrm{o}}} = -\frac{s_{\mathrm{i}}}{s_{\mathrm{o}}} = -\frac{x_{\mathrm{i}}}{f} = -\frac{f}{x_{\mathrm{o}}}, \tag{2.7}$$

where the sign of M_{T} depends on whether the image is on the same side of the optical axis ($M_{\mathrm{T}} > 0$) or not ($M_{\mathrm{T}} < 0$). The magnitude of M_{T} depends only on the focal length and object distance (in the paraxial regime with an ideal lens). With a similar concept in mind the differential rate of change in object and image distances is defined as the longitudinal magnification [4]:

$$M_{\mathrm{L}} = \frac{dx_{\mathrm{i}}}{dx_{\mathrm{o}}} = -\frac{f^2}{x_o^2} = -M_{\mathrm{T}}^2. \tag{2.8}$$

Equation 2.8 states that if one of the magnification variables changes, the other one has to change too. Also, only the first equivalence of equations 2.7 and 2.8 is universal, the latter are valid only for paraxial approximation.

2.3 Lens Systems

Optical systems with multiple lenses, which also cannot be approximated thin anymore, are obviously more complicated than systems based on a single thin lens. More universal definitions of the adduced optical abstractions are thus required. Conveniently, many equations given in Section 2.2 can be applied to lens systems also (within given approximations) just by defining some variables again.

Optical axis defines the route of light where refraction does not happen. Likewise, the spatial input where the light moves in the same direction where it entered throughout the system (if mirrors, diffraction gratings or such are not applied) progresses along the optical axis. The definition leads to the fact that optical centers and centers of curvature of spherical lenses are located on the optical axis.

When the lens is thick, the front and back focal lengths of the system can be different. This is due to the possible different curvatures of the surfaces (i.e. different refractive power) and the distance light has to travel within the lens (determining the location and thus refraction at the output surface). The intersection points of the incoming rays that pass the front focal point (in any angle) and their respective outgoing ray (both shown in Figure 2.2) define the first primal plane (theoretically perpendicular to the optical axis),

whose intersection point with the optical axis is called the first primal point H_1. The second primal plane and point H_2 are defined in the same way, but the incoming ray has to come from the image space and intersect the back focal point. It can be derived that

$$|\underline{F_o H_1}| = |\underline{F_i H_2}| = f_{\text{eff}}, \tag{2.9}$$

where f_{eff} is called the effective focal length. [4]

The rays that proceed in the same angle after the lens as they entered the lens intersect the optical center C of the lens. The lines defined by the incoming and outgoing ray intersect the optical axis at the first and second nodal point of the lens, N_1 and N_2 respectively. The vertices of the lens surface that intersect the optical axis, V_1 and V_2 respectively, are used as a reference to these points.

Focal, principal and nodal points share special properties in the system and are called cardinal points [4]. All aforementioned points regarding the thick lens are illustrated in Figure 2.2.

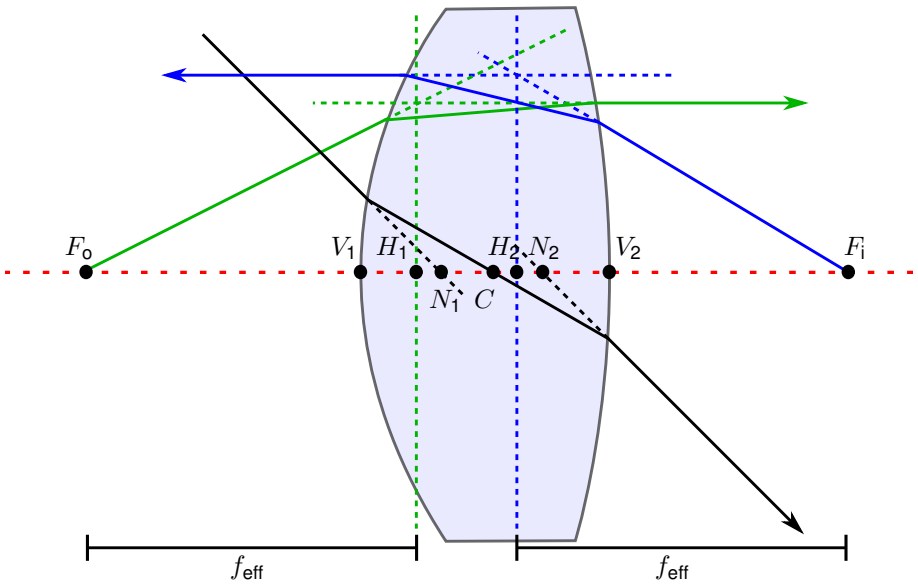


Figure 2.2. *A single thick lens and the rays defining the cardinal points. F_o and F_i are the front and back focal points, V_1 and V_2 the vertices of the lens surfaces, H_1 and H_2 the primal points, N_1 and N_2 the nodal points and C the optical center. f_{eff} is the effective focal length.*

A positive result that can be derived for thick lenses is that the imaging equation, Newton lens formula and transverse magnification equation (equations 2.5, 2.6 and 2.7) remain valid for a thick lens, when the physical dimensions are measured respective to the primal or focal points (H_1 and F_o for object space, H_2 and F_1 for image space) and f_{eff} is used [4]. Other formulas can be derived from these alone, e.g. Equation 2.8.

The aforementioned equations and definitions of cardinal points can be derived also for systems consisting of many lenses, where the outgoing rays leave the whole system, rather than a single lens [4]. This is an important result that can be utilized when considering the properties of optical systems, e.g. the effect of moving a camera sensor that should correspond to the requested image plane. Putting too much emphasis on these methods alone is not enough in real life applications, since ray tracing simulations often provide more substantial value.

Lens systems are often coupled with an appropriate sensor or other apparatus utilizing the imaged information or energy. The term 'focused' is used from now on to describe that the sensor or apparatus of an optical system is at the image plane (or as close to it as practically possible) corresponding to a given object plane.

2.4 Stops and Pupils

Optical systems consist of more than only lenses, such as apertures. Apertures are mechanical, circular holes whose centers are located at the optical axis in rotation symmetric systems. Since lenses are finite physical components themselves, they also serve as apertures alongside their other optical properties.

Stops are apertures with a special property within the optical system. Aperture stop is the aperture or component that determines the intensity response of the system, i.e. how much light (divided as rays) from a given object point can reach the output of the system [4]. This is more influential in the outmost regions of the object plane.

Field stop is the aperture that limits the size of the object to be captured [4]. This means there are object points from which all the emitted rays are blocked by the aperture. There is some given distance from the optical axis from which the objects are blocked and in rotation symmetric systems this is uniform around the axis.

Both the stops discussed here need to be determined for each system separately. Drawing ray tracing diagrams helps to visualize this. In practice, rays are absorbed, reflected and scattered at the physical boundaries of these stops, including lens surfaces. In real life systems, some of these rays could still reach the output of the system, deteriorating the image quality. These rays are called stray light, and the definition includes all types of light rays that are traversing the optics through an unintended route.

Another utilitarian piece of terminology is the pupil. Pupil is the image of the aperture stop, i.e. the aperture stop is now conveyed as the object imaged by the optics that precede it in the system with respect to image space. Pupil can be an entrance or exit pupil, depending on whether the stop is imaged to the image or object space. [4]

Chief ray is the ray of an object point that intersects the center of the aperture stop. It is used in the analysis and simplification of aberrations and correcting them. [4]

2.5 Diffraction

As stated earlier, light can also be handled with wave mechanics. According to the Huygens-Fresnel principle, every point (where scattering, reflection or absorption does not happen) of a wavefront at any time is a source of spherical secondary wavelets [2]. The overall amplitude of the wave at any given point is the superposition of every single wavelet created at some point, considering their individual phase and amplitude. Some points of the wavefront may be obstructed, or the wavefront can be restricted with respect to its cross-section (as in laser beams), leading to wavefront reform, called diffraction.

Diffraction is often divided into two categories depending on how they should be handled mathematically: Fresnel (or near-field) and Fraunhofer (or far-field) diffraction. Fresnel diffraction takes into account that light sources make spherical wavefronts (diverging or converging, depends on the scenario). [2] This applies additional phase and amplitude modifications to each point at the aperture. This is because routes of light are different in length; the phase of the light changes throughout space and the diverging wavelets decrease intensity as the wavefront proceeds. These effects are minimal if the source and destination points are far from the aperture, since the waves arriving and leaving will be planar. This leads to Fraunhofer diffraction, an approximation in real life by definition. [4]

A following rule-of-thumb [4] exists for ratifying the usage of Fraunhofer diffraction:

$$R_{\text{min}} > a^2/\lambda, \tag{2.10}$$

where R_{min} is the distance from the aperture to source or destination, depending on which is shorter, a is the greatest width of the aperture. When the aperture is circular and the diffraction fulfills the Fraunhofer criterion, the intensity diffraction pattern at the destination screen is of the form of Equation 2.11:

$$I(\theta) = I(0)\left[\frac{2J_1(ka\sin\theta)}{ka\sin\theta}\right]^2, \tag{2.11}$$

where J_1 is a Bessel function of the first order [4] and $k = 2\pi/\lambda$ is the wave number. The choice of a circular aperture is worthwhile since many optical systems are rotation symmetric. Equation 2.11 represents an oscillatory function where the amplitude of the first maxima (located at $\theta = 0$) is immensely higher than extrema succeeding it. This means that when the effects of diffraction are considered, only the area within the first minimum is of importance in practice. This area is called the Airy disk and its radius follows Equation 2.12

$$q_1 = 1.22\frac{R\lambda}{a}, \tag{2.12}$$

where R is now the distance from the aperture to the reference target plane. [4] Airy disk represents the fact that point-like object points will never be truly points when imaged to the image plane. This also applies to diffraction in general, e.g. when light is partially obstructed by dust particles. Instead, they appear as dots, deteriorating the image. Optical systems can be said to be naturally astigmatic due to diffraction. Whether this matters or not is dependent on the application.

3 IMAGING SYSTEM PARAMETERS

In this chapter, essential properties of imaging systems are reviewed. Object points are considered as optical signals and the output (images) of optical systems receiving these signals are modeled.

3.1 Field of View and Resolution

In Chapter 2 the concept of maximal object size was discussed. The object space is often wanted to be viewed in the angular space instead of the euclidean one for ease of analysis. Field of view (FOV) is the angle from which the object of maximal size is seen around the axis (i.e. the opening angle) [4].

When using the term FOV the reference point where the opening angle is seen from must be defined. In the euclidean space, the maximal size of the object is defined with the object distance and height, with the front primal point as the reference. For FOV the reference is the entrance pupil [5].

FOV is often designed with camera sensor dimensions and properties in mind. Other aspects could affect the required or optimal FOV. Stray light and optical nonidealities can limit imaging properties tremendously outside of the optical axis. If any of these phenomena become a limiting factor at higher viewing angles (fields), the FOV is often deliberately reduced.

Resolution is defined as the angular size of an image of a photosensor imaged from the image space to the object space [6]. This can also be interpreted as the smallest angular size of objects that can be imaged and collected. The term resolution is not used the same way in optics and consumer electronics. The absolute imaging capabilities depend on many limitations of the system and nature itself (i.e. diffraction). A constant resolution over the sensor (ideally the image plane), consisting of multiple photoreceptors should not be taken as a given.

3.2 Chromatic Aberration

Aberrations are defined as the departures of rays from those determined by Gaussian optics. Aberrations appear in the mathematics when additional Taylor terms (up to third

order) are considered in paraxial rays [4]. Generally, the points of main interest are those that intersect the image plane (how the image is altered), but intermediate analysis can be vital as well (especially when investigating the system).

Aberrations for monochromatic light are called geometrical aberrations. Geometrical aberrations are divided into two subgroups: those that degrade the image (image points become dots or other finite shapes instead of points) and those that deform it (location of the image point changes in 3D or 2D space) [4]. The reasons for geometrical aberrations vary from nonparallel rays to nonideal optical components (finite thicknesses, diameters and varying surface curvatures of lenses).

Chromatic aberrations are aberrations that result from the fact that the refractive index is dependent on wavelength due to difference in light-material interaction properties. Chromatic aberration can be related to shifting of focal points: focal point can shift along the axis or the image plane, named axial and lateral aberration respectively [4]. If the focal length is a function of wavelength, only one wavelength can be at focus in a given image plane at a time. Compensations can be made with achromatized components so that respective images appear at the same image plane (but possibly deformed) [4].

Chromatic aberrations can basically take the form of any type of aberration, but its effect varies as a function of wavelength. Unlike geometrical aberrations, chromatic aberrations only degrade the overall image when multiple wavelengths are present or when the bandwidth of a given wavelength is wide. This is unless different wavelength channels are physically separated somehow.

3.3 Contrast and Modulation Transfer Function

One of the most important properties of optical imaging systems is how detailed information they can provide. Several parameters quantify the sharpness of imaging. The goal of the system defines the primary values and variables that describe performance best.

Depth of field is the lens' property that describes how well the lens can uphold the desired image quality when the object is moved along the optical axis. In practice, this means that as the imaged target is moved or misaligned from the object plane, the image becomes defocused. On the other hand, depth of focus is the image space variant of the spatial defocus phenomena: how much movement and tilt is accepted for the camera sensor respect to image plane while still maintaining good enough image quality. [7]

Contrast describes the difference between lit and dark areas of the image, defined in Equation 3.1:

$$\text{Contrast} = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}}, \tag{3.1}$$

where I_{max} is the intensity (or pixel value) of the most lit area and I_{min} is that of the

darkest area [8]. Contrast alone does not provide sufficient information about the system, since what is to be compared has to be well defined. Alternation of contrast in different scenarios is utilized for example to analyze imaging properties of the system.

Another important variable that needs to be laid out and described is spatial frequency. Spatial frequency defines the rate of change in spatial (periodic) optical patterns. These are demonstrated in the upper parts of subfigures in Figure 3.1 (but due to ease of apprehension these are presented as discrete stripes).

One of the most usual patterns employed in optics is a periodic square pattern. A periodic square (intensity) pattern can be considered as a series of cosine patterns:

$$I(x) = \frac{1}{2} + \frac{2}{\pi}\left(\cos \mathbf{k}x - \frac{1}{3}\cos 3\mathbf{k}x + \frac{1}{5}\cos 5\mathbf{k}x - ...\right), \tag{3.2}$$

where $\mathbf{k}$ is the spatial frequency of the first order term. [4] The inverse of the spatial frequency is called the spatial period (notated as κ), and as per intuition describes the spatial length of one period.

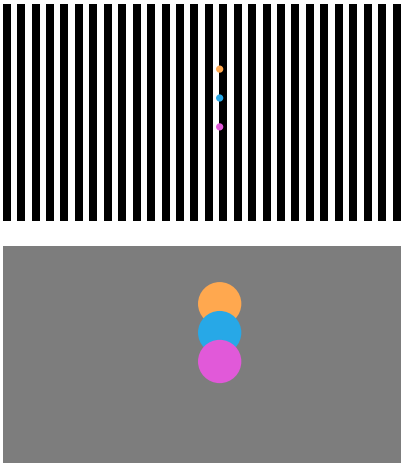
For a practical comparison, we use an analogy from electrical engineering: electrical square pulses in circuits are formed by superpositions of several sine/cosine pulses. When inspecting the pulses with an oscilloscope, the square pulse looks more square-like the more individual pulses from the series in Equation 3.2 are present in the superposition [9]. The main difference between these two scenarios is that the optical one is viewed in spatial domain while the electrical case is viewed in temporal domain (in a pedagogical sense at least, the electronic pulse of course has spatial dimensions and the optical pattern could also vary in time).

Although the periodic square pattern actually consists of an infinite amount of frequencies, it can also be approximated to be a pattern of a single frequency, a sine or cosine of $\mathbf{k}x$. This is beneficial only in scenarios where the central area of multiple stripes is that of interest. With this approximation, the edges shouldn't even exist, but rather be a continuous alternation of intensity (sine or cosine).

As stated earlier, a key factor of an optical system is in many cases its ability to transfer detailed information: the user wants to capture as small spatial features as possible with a given accuracy. Here accuracy mainly refers to the decrease of contrast as $\mathbf{k}$ increases.

Imaging point sources results in practice to finite shapes in the image plane (astigmatism due to diffraction and nonideal lens systems). Thus, in reality, any point at the image plane consists of the sum of intensities of image shapes that suffuse to that point. This results in that the image becomes deteriorated; the light rays are partially mixed up. This leads to decrease of contrast and is demonstrated in Figure 3.1.

(a) *Low spatial frequency* ***(b)*** *High spatial frequency*



(c) *Very high spatial frequency*

Figure 3.1. *Schematic imaging scenarios with different spatial frequencies. Each point in the black and white line strips is imaged by the system, but nonidealities (diffraction, aberrations) deteriorate the image points by modifying them into finite shapes (usually dots). When object points are close, these dots are partially upon each other, resulting in a decrease in contrast. Upper pattern and colored points represent the object, lower ones represent the image.*

Spatial frequency dependency of imaging is quantified with the system's modulation transfer function (MTF). MTF is the relative decrease in contrast as the function of spatial frequency, described by Equation 3.3

$$\mathrm{MTF}(\mathbf{k}) = \frac{\mathrm{Contrast}(\mathbf{k})}{\mathrm{Contrast}(0)} \tag{3.3}$$

The visual representation of the MTF is that when imaging consecutive black and white stripes (zero and high intensity respectively) with same spatial periods, the black strips lighten up and white strips dim as the spatial frequency becomes higher, as shown in Figure 3.1. The most important aspect of MTF graphs is that where the contrast has decreased to a certain percentage. One also must consider how fast these changes happen (i.e. what is the derivative of the MTF throughout frequency space).

When MTF is combined with the Phase transfer function (PTF) as shown in Equation 3.4, the result is the Optical transfer function (OTF) [10]:

$$\text{OTF}(\mathbf{k}) = \text{MTF}(\mathbf{k})e^{-i\text{PTF}(\mathbf{k})}. \tag{3.4}$$

PTF causes phase differences to the wave elements of a superposition pattern (e.g. the square pattern). This results in that the sum of the waves becomes distorted (e.g. the square pattern becomes a nonideal square-like pattern). This phenomenon is called phase distortion.

PTF is especially helpful when investigating the existence of different kinds of aberrations in the system [10]. Aberration analysis based on PTF most likely requires immensely accurate OTF results. As such, PTF and OTF can be omitted in many usual scenarios.

With high-grade optics, the spatial dimensions of the camera cell electronics may become the new limit to the imaging system. The system cannot reach details that would require smaller pixels of the camera cell (which is ideally the image plane). Nyquist frequency $\mathbf{k}_{\text{Nyquist}}$ is defined as half the sampling rate [11], i.e. in camera systems half of the count of pixels in a given unit length [12]. Nyquist frequency is the highest spatial frequency in the image space that the system can still ideally capture; when images of several small spatial details are located at the same pixel some data will be lost.

The phase dependency of the rays approaching the sensor can dampen the signal within a square-like area. This is why a Kell factor (≈ 0.7) is often applied to the Nyquist frequency [12], resulting in Equation 3.5:

$$\mathbf{k}_{\text{Nyquist}} = \frac{1}{2} \cdot \text{Kell factor} \cdot \text{Sampling frequency} \approx \frac{0.35}{\text{Pixel edge}}, \tag{3.5}$$

where the sensor is assumed to be completely filled with sensors, i.e. when pixel edge is reached another pixel immediately appears. Image processing techniques for extrapolating high resolution images beyond Nyquist frequency do exist [12], but are not touched upon further in this thesis.

Even if the camera pixels would be small, and thus Nyquist frequency high, the system will, in the end, reach the diffraction limit. This is called the Rayleigh criterion and it states that two signals must be separated at least with a distance corresponding to the separation between the maximum and the minimum of the signal [3]. These extrema are typically considered to be at the center of the Airy disk and at its perimeter (maximum and minimum respectively) as described by Equation 2.12. This criterion is dependent on the wavelength of light, as seen in the same equation.

3.4 Geometric Distortion

In all imaging setups, the image has a transverse size as established in Chapter 2. In the paraxial regime transverse magnification M_{T} is constant for objects at a specific object plane. As a matter of fact, M_{T} is not constant even for veritably planar object planes in most optical systems. Simply put, the transverse magnification changes across the plane normal to axis, making captured views look streched and/or squeezed.

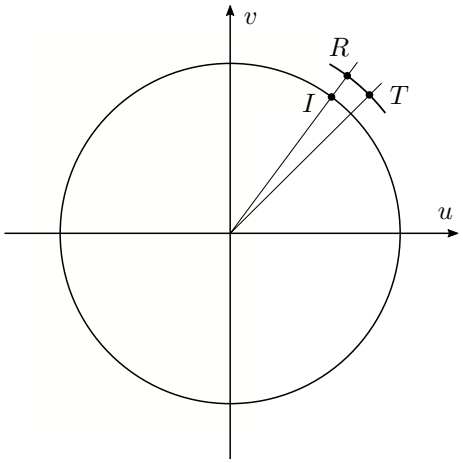
The location of the image at the image plane (ideally this is also the camera image sensor) varies from the ideal point depending on the object's real world coordinates. This phenomenon is called geometric distortion (from now on just distortion), and it is one of the five primary types of geometrical aberrations. Distortion deforms the image in 2D space, while keeping it sharp. [2]

Distortion is divided fundamentally into two categories: radial and tangential distortion. Radial distortion defines the transformation in image space (that is defined by the axes u and v as shown in Figure 3.2) parallel to the radius of the image circle, center of curvature being the optical axis and the radius being the image height. Tangential distortion on the other hand causes perpendicular dislocation of the image respect to the radius, i.e. the dislocation happens over the circumference of the circle when the distortion is small. This is demonstrated in Figure 3.2a.

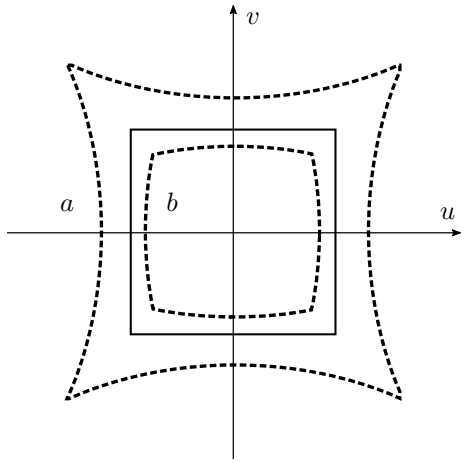
Radial distortion is in most cases rotation symmetric. This means that the transverse magnification of the image is the same for a given radius in the image space. When radial distortion increases M_{T}, the distortion is called pincushion distortion and when it reduces M_{T} it is called barrel distortion [4]. This is visualized in Figure 3.2b. The effect is more impactful for image points further from the optical axis. This is unless multiple sources of distortion are stacked, resulting in e.g. mustache distortion, where M_{T} reaches the same value as in the optical axis at some radius. Radial distortion is mostly caused by erroneous (often aspheric) curvatures of lenses [13].

Tangential distortion is, just like radial, more prominent the further the object is from the optical axis. Although tangential distortion was defined as the dislocation perpendicular to the image radii, it bares neither clear nor straight significance physically. Lens (and sensor) decentering and tilt (in respect to the optical axis) are the major causes of tangential distortion, but they also cause radial distortion. This is because a shifted (decentered or tilted) lens' surface appears aspherical when viewed parallel to the optical axis [14].

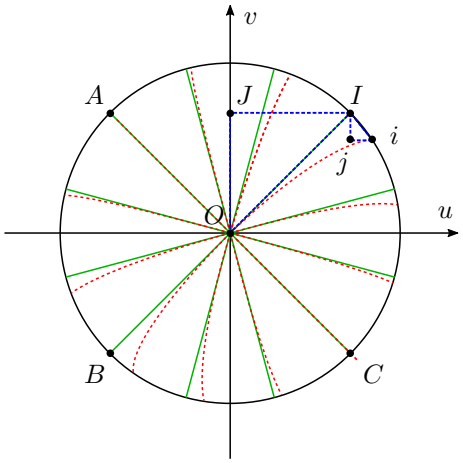
Tangential distortion is in practice (e.g. in machine vision) referred as the movement of image points to a given direction. This is visualized in Figure 3.2c. For some image points the distortion caused by lens shift is purely tangential, since the dislocation happens to the tangential direction. These points define the axis of maximum tangential distortion [13]. The distortion dislocation is perpendicular to this axis.



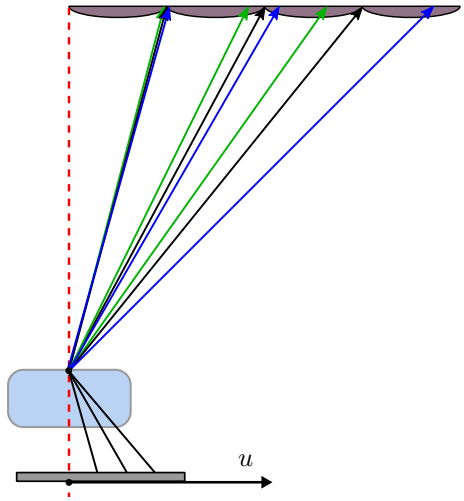
(a) *Showcase of how distortion is defined and divided into two categories. I is the distortionless image point, R with radial distortion and T with also tangential distortion.*



(b) *Radial distortion and its two main subcategories, pincushion (a, M_T larger than distortionless) and barrel (b, M_T smaller than distortionless) distortion. Solid square is the distortionless image.*



(c) *Tangential distortion (red dashed lines) of radial lines (green). O is the image center, points B and I (A and C) define the axis of maximum (minimum) tangential distortion. i and j define the maximal tangential distortion in image space coordinates.*



(d) *Example optical system showcasing what object is imaged to given image heights. Distortionless system is depicted with black rays, pincushion distorted with green and barrel distorted with blue. Tangential distortion can cause 3D dislocation of the object rays.*

Figure 3.2. *Figures for modeling of geometric distortion. Figures 3.2a, 3.2b and 3.2c are adapted from article [13]. Axes u and v form the two dimensional image coordinate space.*

In practise, tangential distortion affects the radial distortion of given image points. This is because the dislocation does not happen over the circumference of the image radius

in a pragmatic sense. In Figure 3.2c, the image points that are on the axis of minimal tangential distortion (line AC), are shifted along the image radius. This shift along the radius decreases as the azimuth object angle gets closer to the axis of maximal tangential distortion.

Issues disclosed thus far lead to a problem refered in article [14]: how should distortion be described so that it is as convenient as possible while remaining its physical implications. For example, tangential distortion is often divided into decentering distortion and thin prism distortion (aka lens tilt, causing prism like effects that result in tangential distortion and chromatic aberration).

In this study, the open source computer vision and machine learning software library OpenCV is used. OpenCV's distortion model and algorithms [15] are used, where the distortion model is divided into four parts: radial, tangential and thin prism distortion and sensor tilt. If sensor tilt is neglected (its effect will be taken into account precise enough in other parameters), the distortion is mathematically expressed in OpenCV as:

$$x'' = x' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2p_1 x' y' + p_2(r^2 + 2x'^2) + s_1 r^2 + s_2 r^4 \tag{3.6a}$$

$$y'' = y' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + p_1(r^2 + 2y'^2) + 2p_2 x' y' + s_3 r^2 + s_4 r^4 \tag{3.6b}$$

In Equation 3.6 x'' and y'' are the distortion dislocation in pinhole camera model respective object space, x' and y' are the division of the object plane coordinates (x and y respectively) and object distance, translated to pinhole camera respective coordinates and object distance, $r = \sqrt{x'^2 + y'^2}$ (i.e. tangent of the viewing angle), k_i are the radial distortion coefficients, p_i the tangential distortion coefficients and s_i the thin prism distortion coefficients.

In the software algorithm, dot target images captured by the camera lens in different angles are used to detect ellipses and their central points. These points correspond to the distorted image points. The coefficients are altered in all the scenarios until a best match is found.

In Figure 3.2c the points that define the tangential distortion on points at the axis of maximum tangential distortion are i and j. It is seen from the figure and Equation 3.6 that now $|\underline{ji}| = x''$, $|\underline{Ij}| = y''$, $|\underline{JI}| = x'$ and $|\underline{OJ}| = y'$, if the space coordinate u and v are changed to the pinhole camera model respective axes. From the figure it can be observed that there are two similar triangles, ΔOJI and ΔijI. If coefficients other than p_1 and p_2 are zero (only tangential distortion is considered), this implicates that

$$\frac{x'}{y'} = -\frac{y''}{x''} \tag{3.7}$$

The line of interest $x' = wy'$ is that which fulfils Equation 3.7. Inserting $x' = wy'$ and

tangential distortion terms from Equation 3.6 into the right side of Equation 3.7 and finally moving the polynomial terms into one side of the equation results in the following condition for the axis of maximum tangential distortion:

$$3p_2w^3 + 3p_1w^2 + 3p_2w + 3p_1 = 0. \tag{3.8}$$

Equation 3.8 is proffered as the method for defining the direction where the image points are translated if tangential distortion is present. Analysis of systems can benefit by the knowledge of the direction of tangential distortion. For example, consider the structure of the camera sensor matrix: if the translation direction is (almost) parallel to the pixel diameter, the smallest distortion that is observable is slightly higher what it would be along pixel edges ($\sqrt{2}$ times bigger). Square-like distribution of the pixels is assumed, which is quite valid for real-life sensors.

One way to rationalize the manifestation of radial distortion is to consider a scenario with an aperture and an ideal lens, demonstrated in Figure 3.3. If the aperture is located at the lens, the chief ray of a ray bundle intersects the lens at the optical axis. When the aperture is moved further from the lens and closer to the object, only the rays coming into the lens in given, higher angles can reach the lens. The chief ray now intersects the lens away from the optical axis. The exploit of paraxial optics can be prolonged if these angles are not too large, leading to the increase of absolute distance between the object and the lens.

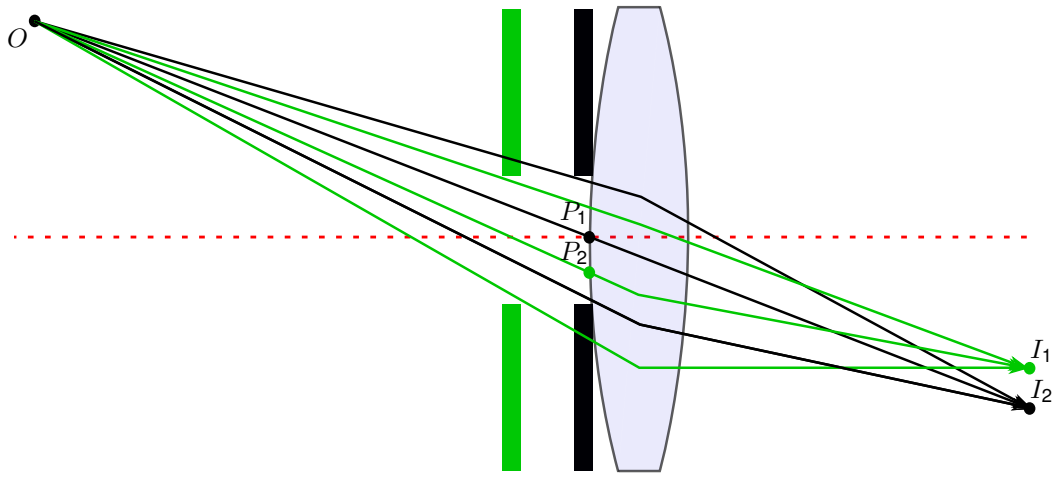


Figure 3.3. *A single lens imaging an object point into an image point with two aperture scenarios. O is the object point, P_1 and P_2 the intersection points of the chief rays with the lens and I_1 and I_2 the image points of the respective scenarios. The magnitude of the vector $\underline{OP_2}$ is greater than that of vector $\underline{OP_1}$, which can be interpreted as elongation of the object distance and thus resulting in a smaller value of M_T.*

Elongation of the absolute light path with an aperture can be considered as an increase of the object distance [4] and reappointment of the optical axis (being now parallel to the chief ray). This leads to decrease of M_T according to Equation 2.7. Thus, barrel distortion is present in the image, as perceivable in Figure 3.3. The on-axis (the real optical axis,

red dashed line in Figure 3.3) objects are not affected by this effect. These objects are located at the center of distortion.

The same way of thinking can be applied in the case of tangential distortion. Changing the lens to a tilted one but retaining some distance between the lens and the aperture (as shown in Figure 3.4) will additionally cause tangential distortion. Two object points opposite to each other in respect to the optical axis have the same object height, but the object distances (now interpreted as the magnitude of the vectors $\underline{O_1P_1}$ and $\underline{O_2P_2}$ in Figure 3.4) are different, yielding the aforementioned causality. Respective optical axes are assumed (which are parallel to the respective chief rays and a slight tilt of the lens is within paraxial approximation).

When the lens tilt lengthens the distance between the object point and the intersection point of the chief ray and the lens (case O_1 in Figure 3.4), M_T becomes smaller according to Equation 2.7. The change is reverse on the opposite side of the optical axis (case O_2 in Figure 3.4).

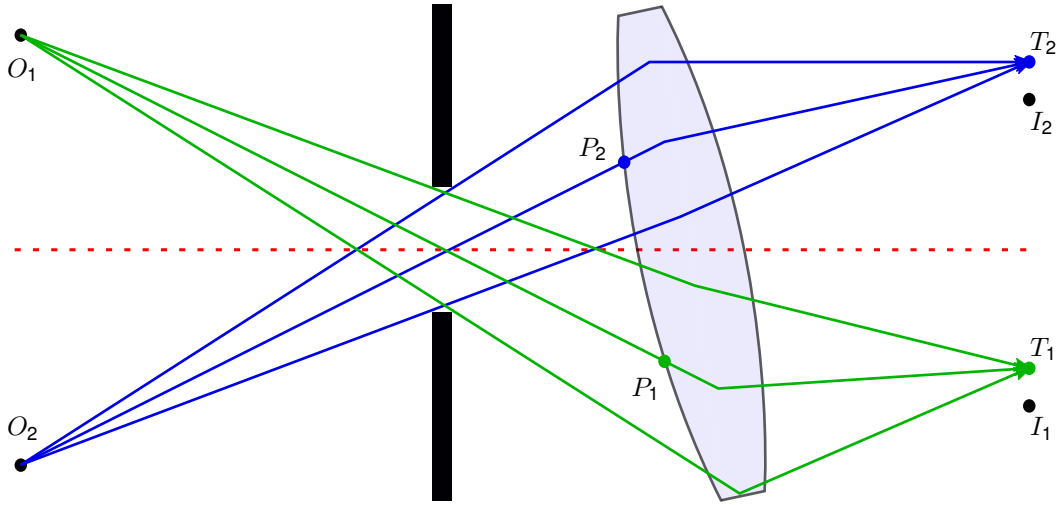


Figure 3.4. *A single lens imaging two object points at the opposite sides of the optical axis. O_1 is the object point with positive y_o, P_1, I_1 and T_1 are the corresponding intersection point of the chief ray and the lens, ideal image point and the tangentially distorted image point respectively. O_2, P_2, I_2 and T_2 are the equivalent points of the case with y_o of opposite sign. The magnitude of the vector $\underline{O_1P_1}$ is greater than what it would be without the lens tilt, which can be interpreted as elongation of the object distance and thus a smaller value of M_T. On the contrary, $\underline{O_2P_2}$ has a smaller magnitude and thus M_T is greater.*

The change shown in Figure 3.4 can also be considered as the overall shift of the image to a given direction (determined by the axis of maximum tangential distortion) in the two-dimensional image plane. The outmost parts of the image are shifted the most, since the effect caused by the lens tilt is greatest there. Objects close to the optical axis are less shifted while the object point on optical axis is not shifted at all.

The way of thinking demonstrated in figures 3.3 and 3.4 can come off as quite baffling, since without the aperture the rays would not focus to a single point (paraxial, stigmatic

case and ideal lens were assumed). Usage of this rationalization is not applied in any real distortion calculation (instead ray tracing is utilized), it is meant only for thought. The valuable content of this argument is that even mechanical tolerance and apertures can cause distortion, not only lenses and their surfaces. This argument can also be exploited in other scenarios, for instance, when apertures are shifted.

It also needs to be brought up that the cause of distortion in figures 3.3 and 3.4 was caused by the alternation of the object distance. In cases where the object distance is high to begin with (e.g. the first lens of a system with a high focus) this should not cause any distortion. However, in optical systems intermediate and virtual images are often used in design. If the distances of these points in respect to the primal plane of optics preceding it are small enough, apertures and nonidealities in the lens and their assembly could cause more distortion.

The distortion model of a system is actually equivalent to the angular point-to-point mapping function of the system. This can be seen and proved based on Equation 3.6, where angular (tangential) object coordinates are translated to image space respective coordinates and mapped on the image plane. However, this angular mapping function does not take into account how the object points are deteriorated when they are imaged to points (MTF etc.).

Distortion is often deemed solely as a nonideality of the optical architecture and the respective manufactured system. Of course, tangential distortion is practically always a hindering trait, while radial distortion not so. Many applications lean on some specific angular response (i.e. angular mapping function) in order to produce or record optical signals of interest (e.g. a case where object plane is not planar). Therefore, correcting distortion in e.g. software calibration can be a faulty way of processing data, depending on the application.

3.5 Flatfield and Dark current

In traditional ray optics the number of rays collected by a system as a function of the object location is often discarded. In pragmatic applications this way of thinking is amiss in almost all real-life systems, since both information and energy can be lost depending on the input. Objects that are far off-axis contribute less rays in the system input. This phenomenon is called vignetting [4] and it is visible in captured images as darkening areas at the edge of the image.

Some sources consider the term vignetting as the highest abstraction level of image darkening for off-axis object points. It is also normal to see vignetting being treated as only one of the causes of darkening alongside effects like cosine-fourth and pupil aberration. Cosine-fourth is the phenomenon that at higher input angles the perpendicular area of the entrance pupil is smaller, resulting in less light exposure to the system [16]. Pupil aberration considers the fact that refraction along the lens surface is not linear and the

apertures supervening the lens might receive a nonlinear light input [17].

Other forms of vignetting can also appear. The nonuniform angular gain of electronic sensor arrays [16] causes darker output at higher input fields. A pragmatic example of this phenomena is shown in the datasheet of ON Semiconductor PYTHON 16 CMOS sensor [18], where gain rapidly decreases to less than 20% when the field reaches 30°. Spherical distribution of light from an object point results in decreasing intensity as distance from the source increases. If apertures are large, the intensity of light emitted from an object point can thus vary over the apertures, resulting in vignetting. Distortion can also change the intensity response at the image plane (barrel distortion collects more light to a single spot from high fields while pincushion distortion spreads the light at high fields).

Vignetting is normally strong only close to the edges but is seen as a descending intensity distribution throughout the image from center to the edges. Not only vignetting is causing variations in the intensity profile of the system. Dust, dirt, scratches etc. are also part of the problem, especially in consumer products (cheap production or user caused damage and dirt).

Aforementioned problems are solved with an image of a very close, out of focus plane with a uniform angular intensity distribution profile (in respect to each point in the plane) captured by the system. It is crucial that the object plane is placed at the front of the system, because only this way it is possible to capture the complete angular input without optical noise (e.g. scattered light). This image is called the flatfield of the system [19] and it describes qualitatively the intensity response of the system.

Captured images are scaled with the flatfield so that brightness of all image points (in practice pixels) correspond to the correct amount of light their respective object emits. A more straightforward way to express this is that the intensities of all points are multiplied by a respective value that would result in the same intensity all over the flatfield image. This results in a more correct image. This process will be explained further in Chapter 6.

In the areas where intensity response is low (such as the outline of the imaging circle) the accuracy of the measurement is also lower. This is due to the fact that when the region of highest intensity response uses the complete measurement range supported by the camera (determined by the bit rate of the camera) so that it does not saturate (i.e. reach the highest supported measurement value), the region of lowest intensity response uses only a small part of the range. For example, with an 8-bit camera the center of the image utilizes the pixel value range of 0-255, while the outmost region of the image can reach only a fraction of it (e.g. between 0-40). The conclusion is that the important aspect of flatfield is in many cases not the causes (vignetting, cosine-fourth, pupil aberration etc.) of dampening but how much the accuracy of the application is lowered. Flatfield does not solve this issue.

Since optical systems nowadays consist also of electronics, the problems in electronics are also carried over into optical systems. The most fundamental parts of the optical electronics are the image sensors. Image sensors collect photons with photodiodes where

photons excite electrons. After the exposure time has passed, the electrons participate in a current that enters an amplifier circuitry. The output of this circuitry is analyzed in an analog-digital converter, resulting in digital pixel values that are linear to the number of excited electrons (and also to the energy of the input light, i.e. number of photons in designed operation). [20]

Image sensors can be faulty, usually seen as hot-pixels where the electric current is too high even without any electro-optic cause. Even in ideal circuits some electric current exists without stimulus by photons. This is called dark current. This is best described when no light at all enters the system, but the digital image still has some kind of unwanted responses. This response is called the dark frame of the system [19].

Dark current can be considered as temporal noise of the system. Sensors are susceptible to other forms of noise too. This includes thermal, read and shot noise. Thermal noise is related to the increase of measured current due to temperature effects. Read noise is related to the current amplification process. [21] Shot noise accounts the fact that the carrier (electrons, holes i.e. charge vacancies) generation is a statistical process. Shot noise is in general a natural physical phenomenon predicted by the quantum mechanics. [22] This means that dark current also has some temporal variance due to randomness [21].

The signal-to-noise ratio (SNR) of the sensor is calculated [21] with Equation 3.9

$$\mathrm{SNR} = \frac{\langle I_{\mathrm{S}}^2(t)\rangle}{\langle i_{\mathrm{dark}}^2(t)\rangle + \langle i_{\mathrm{thermal}}^2(t)\rangle + \langle i_{\mathrm{read}}^2(t)\rangle + \langle i_{\mathrm{shot}}^2(t)\rangle}, \tag{3.9}$$

where $\langle I_{\mathrm{S}}^2(t)\rangle$ is the mean-square of the signal current, and the denominator values are the corresponding values for the individual noise currents. When the signal current equals the saturation current, i.e. the maximum value that can be registered, the SNR value represents the dynamic range of the system [22].

In the case of systems where optical signals are quite low, thermal noise dominates [21]. In this situation all other forms of noise in Equation 3.9 can be neglected. However, for convenience, dark current can also be included, since the dark frames measured in operating temperatures include both of these noise types.

The temporal wavering of the noise can be taken into account by averaging methods. However, noise can also vary spatially between the separated pixels. This is why the signal and noise components in Equation 3.9 should be considered pixel-specific. The same equation can be used for the whole sensor, but the overall noise level has to be determined from the data. One way to do this is to take the average noise and describe the error margin of the SNR with the standard deviation of the noise (e.g. as units of electric charge) [22].

3.6 Optical Axis and Rotation Symmetry

In reality, the definition of optical axis laid out in Chapter 2 might not provide any real value when evaluating the qualitative properties of a system. Optical components always have some tolerance in their structure and assembly in the system. This troubling fact necessitates that a new definition of the optical axis has to be agreed on in order to quantify it.

Intuition states that the optical axis has to be the center of the system, but the center in respect to what exactly? The most simple and concordant parameters to quantify the location of the optical axis are the interception point coordinates of the axis and the camera sensor. Ergo, the optical axis can be pinpointed in every image captured.

Many candidates can vouch (or at least be part of the verification process) the location of the optical axis. The measured and calculated center of distortion is a valid choice for the optical axis. After all, in wide-angle lenses the radial distortion is often quite formidable in raw images and the center is easy to verify. On the other hand, vignetting should be symmetric (in rotation symmetric systems) over the aperture stop, whose center is located on the optical axis. This means that the center of the intensity distribution in the flatfield could be defined as the optical axis.

Due to aberrations and other phenomena the imaging properties of lens systems are worse at higher fields (aka with larger object sizes). In theory, the imaging capability should be at its best at the optical axis. Measuring MTF values at different angles and doing some interpolation would unveil the location of the optical axis.

In rotation symmetric systems the point in the image sensor that images the same object point when rotating the system (or the point that results in the best fit in such analyses) gives information about the optical axis. The center of mechanical rotation can bring upon valuable information about the quality of the whole system. In many applications though it is definitely not the best way to quantify the axis itself, since nothing guarantees that the mechanics are aligned harmoniously with the optics.

Rotation symmetry means that the optical properties of the system are same at a given radius from the optical axis. Quantifiable limits to the possible tolerance and errors of the optical variables at a given radius should be mandatory in order to distinguish rotationally symmetric systems from those that are not. For example, the values for M_{T} and MTF50 should be allowed to vary only by an agreed, high-fidelity amount at a given radius.

Rotation asymmetry is often associated with stray light, i.e. asymmetry is one type of manifestation of stray light. It is also a quite vexing form of nonideality, since many computational optical models are assumed rotation symmetric (or at least no arbitrary asymmetry is allowed). A good example of this is the distortion model used in this thesis (Equation 3.6). This does not, however, apply in general to ray tracing software applications.

4 NEAR EYE DISPLAY PARAMETERS AND QUALITY ASSURANCE

Firstly, the human eye is reviewed as an imaging system in this chapter. This information is connected to the properties of near eye displays, which provide optical signals to the user pupil in the form of virtual images. Virtual images can be considered as objects that are prolonged and enlarged (with a lens) from physical objects, e.g. pixels of a screen.

4.1 Human Eye

The human eye is basically a wide-angle imaging system, engineered by the evolution of mankind. It consists of fewer lenses and other optical elements compared to technological, modern imaging systems. It is substantially a very comprehensive system, having a large FOV and quite homogeneous imaging quality through these fields (i.e. a bit lower quality at the center and higher quality at slightly wider fields compared to man-created imaging systems) alongside the fact that for a normal person these properties are almost mirror symmetric for the eye pair [23]. The eye is not perfect, bearing a lot of aberrations and other nonidealities (like scattering media) that increase by the aging of the person.

In this thesis mainly the optical structure of the eye is of advantageous information, so many biological viewpoints are disregarded. The structure of the eye is complex enough to cause a lot of trouble to researchers fitting an appropriate model that represents the eye [23]. This is due to difference between individuals (genetics, age etc.) and the possibility of altering optical properties in an agile manner, such as pupil size and focus.

The first element of the eye is a meniscus lens (first surface has positive radius of curvature, the second negative) called the cornea and its anterior surface is called the conjunctiva. The cornea can collect light from high input fields, and it has a high refractive power, partially because of the graded refractive index (GRIN) and conic shape of the conjunctiva. The cornea is followed by an aqueous humor and the iris, which is the aperture stop of the eye. Iris size can actively vary, resulting in different exposure to light (causing vignetting and even aberrations to change). With lower exposure the signal-to-noise ratio is smaller due to quantum noise. [23][4]

At the rear side of the iris is a convex lens (whose latter radius of curvature is higher than the first) called the crystalline lens. Crystalline lens is attached to ciliary muscles that can

stretch or loose the lens, changing its curvature and thus refractive power. The property of changing the focus distance dynamically is called accommodation [24]. Diameters of the crystalline lens are much smaller than that of the cornea. The crystalline lens also has a GRIN structure. [23][4]

The crystalline lens is followed by vitreous humor and after that the retina. The retina serves as the optical sensor of the system, photoreceptor cells in the place of electrical photosensors. There are two types of photoreceptor cells, rods and cones, which function in waveguide like manner (and afterwards the events take place in the nervous system of the human). [23][4] The schematic illustration of the human eye is shown in Figure 4.1.

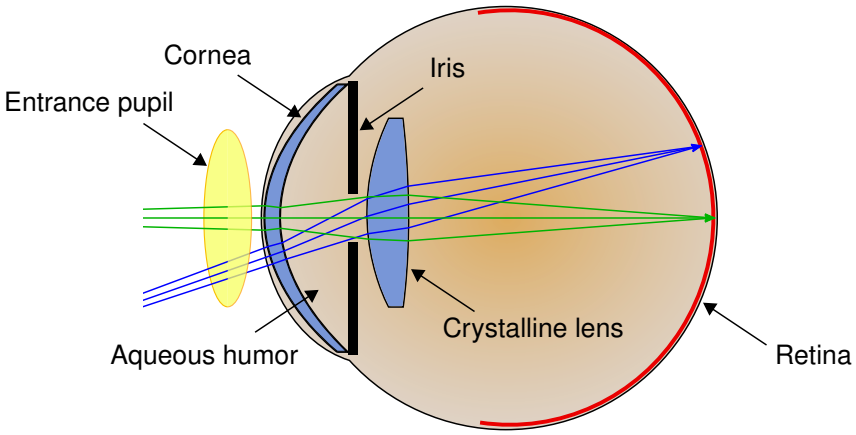


Figure 4.1. *Schematic optics of the human eye with two imaging scenarios: on-axis and off-axis. The real eye has relevant decentering with its components and the pupil is larger and spherical [23], allowing higher input fields.*

Human vision enables color detection, which is based on the cone photoreceptor cells. There are three main types of cones, L, M and S cones, referring to their responsivity distribution: L being centered around a long wavelength (red), M for middle (green) and S for short (blue) [23]. Since humans can perceive more colors than these three main types, it must mean that the colors whose wavelengths are in between the distribution centers are formed as a stimulus combination of all three types of cells. This is the foundation of the Red Green Blue model (RGB) used in electronic color displays, where only three types of light sources are mainly used, packed so close that they seem like a single point source.

The quite simplistic optics of the eye result in a nonplanar image plane, which is compensated by the curvature of the retina (so that all light is focused on the same surface). The image has a negative transverse size, implying that the image is flipped, which is fixed by the human brain. Many other properties (such as aberrations, caused partially by the fact that there is notable decentering with the components) are also rectified by the psycho-physical phenomena. Some effects become apparent after a long time, near the elderly age, and are not compensated (enough) by the brain. [23][4]

The entrance pupil is likewise an ellipsoid surface. This is more dominant further from the optical axis, in order to cover high viewing angles. The true shape of the entrance pupil varies from person to person and models of it are based on empirical data. [25]

The imaging quality of the eye is degraded also by scattering and diffraction, mainly because of the heterogeneous structure of the media, caused by suture lines, damaged tissue etc. Thus, the human eye is far from being diffraction limited. It is also not rotation symmetric merely due to shifted components. Likewise, the vertical FOV is lower compared to the horizontal FOV [6].

4.2 Display Technologies

Near eye displays (NED) and Head mounted displays (HMD) are often mixed with each other. In this thesis an intuitive distinction is made between these two. NED is the technology of providing virtual images into the object space of the user's eye (or rather pupil), e.g. a single screen and its optics form a NED. HMDs are the products that usually consist of two NEDs (one for each eye) and other technology in order to fullfill the purpose of the device (i.e. alter user reality).

Current HMDs use different kinds of screen technologies. Usual virtual reality (VR) screens are made of liquid crystal matrix displays (LCD) or organic light emitting diodes (OLED) [26]. Between the human eye and the screen are the optics. All HMDs have to make the image of the screen virtual, that is to say, use virtual image points as object points for the human eye. This way the screen seems further and bigger compared to the real physical dimensions.

Augmented reality (AR), or sometimes referred as mixed reality (MR) (especially by Microsoft), displays have a more diverse catalog of display technologies used. For example, the Microsoft's Hololens uses a liquid crystal on silicon (LCOS) as the light source and directs it to the right output with waveguides and diffractive gratings (lenses etc. are also used). In order to achieve uniform light intensity at all output angles of the display, the diffractive grating must have a nonlinear output (i.e. relative output increases with distance travelled in the waveguide). [27] Combinations of AR and VR also exist, but in these devices the reality which is augmented is simply the footage of the cameras placed at the front of the device [28].

Whatever technology is behind the HMD display, they share a common trait: they only have to use three wavelengths in order to cover the complete color stimulus region of a human (thanks to the RGB model). For VR the most helpful aspect of the RGB model is that chromatic aberration (mainly difference in the radial distortions between wavelengths) is a bit easier to fix software-wise. In the case of AR, consider the Hololens: waveguides and diffractive gratings have to be designed only for three wavelengths.

All display technologies strive for the ultimate goal: making the FOV of the display larger,

until it reaches the FOV of the human eye. The physical limitations (aperture dimensions etc.) reduce the FOV of the system. Sometimes this is desired, but in HMD applications this is surely not the case.

4.3 Imaging Parameters in Near Eye Displays

The optical concept of NEDs is somewhat similar to imaging systems like cameras: they form image points out of object points. However, as mentioned afore, the image points are virtual and $|s_\mathrm{i}| >> |s_\mathrm{o}|$, meaning that the screen appears to be much further and larger according to equations 2.7 and 2.8. These parameters are used when defining the FOV of a NED; the difference between FOV of a NED and that of an imaging system is that with the former the maximum angle is seen from the exit pupil in the image space instead of the entrance pupil in the object space.

Aberrations (such as geometrical distortion) and MTF are present in NEDs too. The meaning of these parameters has to be interpreted as the quality and properties of the virtual images instead of real images. Therefore these parameters were more straightforward to depict in Chapter 3.

The total geometrical distortion and MTF of a system are a combination of those of its components, such as separate lenses. This also applies in a setup composed of a NED and an imaging system (e.g. human eye): the total distortion or MTF (or any parameter) is a gestalt of the two subsystems. The total MTF is easy to calculate, it is simply the multiplication of the MTFs of all the components of a system. This is because MTF is simply the relative change of contrast as a function of spatial frequency, e.g. if two components decrease a given frequency by 50% each, the total relative contrast is 25%.

The flatfield, optical axis etc. of a NED can also be characterized, although the latter should be defined first. This is because many NEDs are not rotation symmetric due to the asymmetry of the human eye. The flatfield can be compensated simply by increasing the intensity at the outer sides of the screen.

4.4 Eyebox

The eyebox is the volume in the image space of a NED where the user can see the whole FOV of the device [6]. The most beneficial way of defining the optical axis of a NED is by utilizing the center of its eyebox (the center should be on the axis). The eyebox volume is centered around the exit pupil of the system. The brightness of the NED decreases if the volume of the eyebox is increased. This is because the total distributed energy does not change (unless the power of the light source is changed).

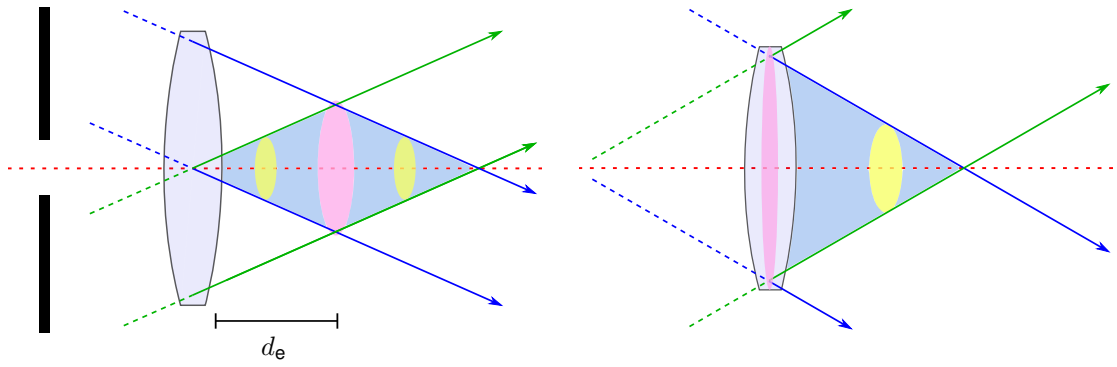
The distance between the last optical element and the exit pupil is referred as the eye relief d_e [6]. Humans have two eyes, and the relative distance between the pupils is

called interpupillary distance d_{i}. With these three metrics introduced afore the optimal use case can be derived.

It is simple to justify the nature of the eyebox. Consider a large view of an area and then place a thin wall with a circular hole in it in front of you. You must place your eye exactly at the hole in order to see the whole view. The exit pupil acts just like this hole, since no light can come out of the system outside of the pupil.

If both the exit pupil of the NED and the entrance pupil of the eye are assumed to be circles (i.e. rotation symmetric and aberrationless system is assumed, which is always amiss for the latter) the eyebox is actually a volume of two cones with the same base (which is the exit pupil) and magnitude of height, as shown in figure 4.2a. The most ideal scenario is that the centers of the pupils are at the same place. When the user pupil is partially outside the eyebox, vignetting occurs (even if the pupil is inside the eyebox vignetting can happen due to nonideal intensity distribution of the system).

For simplicity, the angles a certain object is seen at are assumed to be the same all over the exit pupil (i.e. it is assumed that the object is at infinity, which is justified since exit pupil sizes are small compared to the virtual image distances). This scenario is illustrated in Figure 4.2



(a) *Eyebox of pupillary optics* ***(b)*** *Eyebox of nonpupillary optics*

Figure 4.2. *The two traditional eyebox (painted blue area between the output rays) scenarios. In the pupillary case an aperture stop (black surface on the left with a hole) forms the pupil (pink hole) at eye relief d_{e}. In the nonpupillary case the exit pupil is actually the lens element itself. Yellow holes represent the user entrance pupil. Theoretically most HMD devices have a nonpupillary eyebox, but due to nonidealities the shape is most likely different.*

Many HMDs utilize optics that are quite similar to magnifying glasses, i.e. the exit pupil is actually the lens itself (the lens is the aperture stop and the last optical element and thus the exit pupil) as in Figure 4.2b. This kind of optics are called nonpupillary, although quite misleadingly since all real-life optics have some limit on how much light can be collected, i.e. there is always a pupil. Unlike in Figure 4.2a, the eyebox is strictly close to the optics. Nonpupillary implementation is used in many VR devices, especially in budget models.

The eye relief of nonpupillary optics is, by definition, zero. HMD optics contain nonidealities and design choices that can limit the eyebox size. Light rays that are emitted by the screen in high angles are an example of this.

Consider a Fresnel lens (used in HMDs due to light weight etc.) that consists of several slices of sphere surfaces, separated by rough edges. Light rays that are deflected by edges at the opposite side of the optical axis turn into stray light. Patents by Google [29] and Oculus [30] show that this is indeed possible, the latter is even acknowledging the noxius nature of this straylight. LCD screens (e.g. a computer monitor) emit less light in higher viewing angles, and this happens with HMDs also. The deduction is that eyeboxes that should ideally be that of Figure 4.2b resemble in practice that shown in Figure 4.2a.

Pragmatically, a sweet spot with full FOV and minimal vignetting and aberrations is always found from a finite distance in respect to the surface of the last optical element of the system. Thus a definition of an effective eye relief is justified. In this thesis, effective eye relief is defined as the perpendicular distance of this sweet spot in respect to the surface of the last optical element.

As can be seen from Figure 4.2, there is a trade-off between the FOV and the eyebox and eye relief. If the outmost fields were at even higher angles, the eyebox would be smaller and closer to the optics. This is one of the crucial challenges of HMD design.

In many cases the effect of aberrations, especially tangential distortion, is dismissed when considering the eyebox. Even inside the eyebox notable aberrations can occur. The most general example of this is people whose interpupillary distance is not the same as that designed in the HMD (especially children, whose interpupillary distance is small). Another cause for the inadaptability of the HMD to the user is an undesired location of the eye rotation point, causing varying pupil mismatch with different viewing angles.

As was implied before, the exit pupil of the NED does not need to be rotation symmetric. Instead the eyebox could very well be stretched or otherwise asymmetric. In practice the most crucial thing is to agree on some rule how to distinguish when the eyebox is outstripped, i.e. what is the reference eye pupil size and highest reference image/pixel intensity value that is still accounted for eyebox (e.g. 95% of maximum value).

4.5 Vergence and Eye Tracking

The property of looking at the same object with both eyes and forming a stereoscopic perspective of it is called vergence [24]. This is the foundation of 3D perception. HMDs try to replicate a similar environment in their design.

Vergence should not be confused with accommodation (focus of the eye). In real life situations these two properties work in tandem: the focus and stereoscopic distances are equal. However, in the current commercial HMD devices the virtual image distance is fixed while stereoscopic view can be altered. This is called the vergence-accommodation

conflict (VAC) [24] and remains a tedious problem in present-day HMD devices, causing nausea to the user.

The only natural way to alleviate the effects of VAC is to introduce several, logarithmically separated focus planes, since the crystal lens refractive power has to change denser the closer the object is. This is an arduous task, which would require novel technology, such as light-field displays [24] (based on, for example, many transparent screens like that of PlayNitride's [31], shown lastly at DisplayWeek 2019). Diffraction could still cause severe problems in solutions of this kind. It is also about tradeoff between properties that make the views realistic and natural (e.g. resolution versus number of focus planes).

Another, technology-wise more feasible solution is eye tracking [24]. The direction of the eye pupils can be distinguished and the interception point of the optical axes of the eyes can be calculated. The interception point forms an opening angle between the eyes that can be tracked. The object of interest at the virtual digital image can be deduced and the appropriate focus plane can be swapped in. This kind of hardware requires detector cameras and a moving screen and/or optical elements. A more trivial, but debatable, solution is to cause software based blur of the objects.

These topics related to properties that are yet to be implemented in commercial HMDs are important, because the characterization optics (the HMD Eye or its possible successor, product/system family etc.) have to take these into account in the long run. Technology industry can be hard to predict, but the fundamentals of physics do not change.

4.6 Defects and Factory Environments

The screens of a NED can contain variety of defects. Defects can be mechanical, chemical, electrical and so on. The most usual defects depend on the type of the screen alongside the adequacy of the factory environment. The defects of NEDs are more crucial than those of traditional displays, because when the virtual image is formed, the defects are also magnified and distorted. In small devices like smartphones and HMDs no major screen defects are allowed at all in general. [14]

LCD screens are often tainted with Mura defects. Mura defects are dark or bright lines, spots or other regions in the screen where the liquid crystal material is non-uniformly distributed or mixed with foreign particles. [32]

The OLED defects are caused mainly by electronical faults: open and short circuits (OLED cathode, anode and driver connections) [33]. Semiconductor material (substrate, epitaxially grown device) related issues also cause defects [34].

Since the NED also consists of the optics, the defects in these elements have also to be inspected. The number of optical elements is often scarce in HMDs, since many optical nonidealities, such as aberrations, can be fixed with software. Regarding distortion, HMDs usually bear only radial distortion and many HMD standard development

kits (SDKs) consider only this. The reason is that the system consists of minimal optics that are manufactured and assembled in a very controlled environment and are thus less prone to error in terms of decentering, dislocation and angular shift. Thus, the manufacture is easy and cheaper than optically advanced systems and any appearing distortion is rather easy to correct.

The operation procedure of HMD factory testing depends on the type of the device. Low end product (i.e. cheap HMD devices) testing is usually done after the assembly of the device. Risks of faulty products are higher, but the total manufacture costs of a single HMD are really small. Throwing away bad products becomes cheaper than multiple inspections during several phases of manufacture. High end products (especially their lenses) go through repetitive manual inspection at clean room environments before assembly and integrated tests.

In general, the testing principles after assembly are quite similar between low and high end HMDs. Since low end products require a more holistic inspection process after assembly, new challenges are faced. The main concern with cheap HMDs is that can screen and lens defects be detected through the user interface of the device. This depends on the resolution (or rather Nyquist frequency), noise levels, dynamic range and MTF of the characterization optics and the HMD itself. It is then evident that defect detection becomes more difficult the further light sources are from the optical axis.

5 HMD EYE LENS AND CAMERA

In order to measure the optical signals users would perceive from near eye displays, a devoted optical measurement system is necessary. This is the foundation of the HMD Eye. The last section is dedicated to comparison of similar competing systems.

5.1 Design Principles

The main focus of this thesis is OptoFidelity's HMD Eye, an optical camera system that mimicks the performance of the human eye. The HMD Eye is shown in Figure 5.1. HMD Eye is a wide-angle lens with a mechanical, external entrance pupil that is of the same size as the human entrance pupil. This is beneficial, because the HMD Eye can be easily used with HMD devices, whose eyebox, illumination etc. is designed for the human eye, while maintaining low noise levels. This allows the measurement of HMD devices from the perspective of the user, i.e. universal analysis possibilities. With a generic fisheye lens this wouldn't be possible.

The analysis procedures are put into practice in OptoFidelity's HMD IQ test station. HMD Eye is used to measure optical signals (virtual images) sent by head mounted displays (HMD). HMDs in the HMD IQ are called devices under test (DUT).

(a) *Side view of the HMD Eye* ***(b)*** *Front view of the HMD Eye*

Figure 5.1. *Images of the HMD Eye.*

Electronic cameras contain a planar image sensor, whereas the human eye has an ellipsoid (almost spherical) photoreceptor surface. This demands much more functionality

from the optical elements of the HMD Eye. There are many times more components within the HMD Eye compared to the human eye (although the latter has the perk of GRIN structures).

HMD Eye contains a very high SNR camera. The choice of the camera reflects to other properties as well. The camera sensor size determines the FOV, while the pixel density determines the resolution. In real life there is a trade-off between these parameters, if financial costs are heeded.

The camera sensor's gain as a function of input electric field can decrease the dynamic imaging capability of the system. Since gain decreases with higher fields, it is a good practise to design the optical image points so that the rays intercept with only slightly divergent angles respect to the optical axis. This is taken into account within the design of the HMD Eye.

The camera sensor is often covered with a protective glass. This can have a debilitating effect on image quality, since it can cause unwanted internal reflections (amongst other vulnerabilities) nearby the sensor. The protective coating is neglected in the HMD Eye, since high-class imaging quality is demanded and the sensor is well protected after HMD Eye assembly. Likewise, straylight analysis is more legitimate with this kind of configuration, since total internal reflections of sensor must fall below 3%, which is the limit set by ISO-9358 standard [35].

In VR devices a high FOV is often a demand and resolution comes second. In AR the FOV of the device is often low due to technological limitations and the resolution is a keen interest for research purposes. HMD Eye answers both needs by offering two camera options, Allied Vision Prosilica GT4096 [36] and Allied Vision Manta G-1236 [37]. The former has a larger sensor (ON Semiconductor PYTHON 16 CMOS [18]) size, corresponding to 120 degree FOV and 29 pixels-per-degree (ppd) resolution. The sensor of the latter (Sony IMX304 [38]) provides a higher, 41 ppd resolution and approximately 100x75 degree FOV (it is actually a bit higher than this).

One of the goals with the HMD Eye is to reach the human eye polychromatic MTF, that has an approximate MTF50 of 10 cycles per degree (cpd) [39], depending on the size of the human eye pupil. The HMD Eye has much less aberrations than the human eye, apart from radial distortion. Bad quality of the components and their assembly can hinder the imaging quality.

HMD Eye has designed barrel distortion just like pretty much every wide-angle optical system. The radial distortion is linear, meaning that the image point at the sensor is linearly dependent on the polar angle. The HMD Eye is thus ideal for imaging angular space objects. Planar object planes can also be analysed easily with the help of software corrections based on the distortion calibration.

5.2 Competitor Products

In this section, three different competitor products for HMD Eye/IQ will be discussed. Each one has come to a different solution on how to characterize NEDs. Only one solution is similar to the HMD Eye, and could be considered as the true rival product.

Trioptics offers two solutions [40] for VR lens and module testing, ImageMaster Lab VR and ImageMaster PRO VR. Lab is based on a swiveling telescope (+/- 2 diopter adjustment), enabling measurements up to 180° FOV. PRO utilizes up to 27 cameras for a proper measurement range and is compatible with a similar telescope camera setup as the Lab. The emphasis of these systems is high-end MTF measurements.

Gamma Scientific has ended up with a design [41] where a telescope camera system is connected to a spectroradiometer and a robot allowing multiple degrees of freedom. The resolution is 250 ppd, FOV 16x12°, 2-5 mm aperture and +/- 10 diopter adjustment (+/-0.1 m to infinity). Photoluminescence, spectral and defect measurements are emphasized in this solution.

Lastly, Radiant Vision Systems offers a wide-angle measurement camera system [42], resembling the HMD Eye. The FOV of the system is 120°, 3.6 mm aperture at front of the optics, manual focus between 0.25 m and infinity. If Radiant ProMetric Y16 or Y29 is used with this system, ideally resolutions up to 54.8 ppd are achievable with FOV of 120°, assuming that the image circle fills the horizontal and vertical dimensions of the sensor.

HMD Eye is a more sophisticated eye simulating/mimicking system than what Trioptics and Gamma Scientific are offering, but it is not necessarily competing directly with them. These two offer first rate measurements in their own right, but neither is the all-around solution. Having to use several cameras or moving them adds uncertainty to the measurements where higher fields are of interest.

The Radiant Vision System's product on the other hand is frankly the rival product to the HMD Eye. Both are wide-angle, small entrance pupil lens systems, with the purpose of replicating how the DUT is seen by the user. The main differences are that the resolution of the AR/VR Lens is higher than that of HMD Eye whereas the FOV is higher in HMD Eye (if AR/VR Lens is assumed to be attached to one of the cameras given in the respective data sheet). The AR/VR lens might also not be compatible with apertures of other sizes than 3.6 mm, at least according to the datasheet.

In the end, the imaging properties like MTF bear a lot of significance in the end product. The datasheet for AR/VR Lens is not sufficient for this comparison. Higher resolution might not be as promising as it sounds if conditions related to optimum imaging performance are not met.

6 HMD EYE CALIBRATION METHODS

The calibration and characterization methods for HMD Eye are laid out in this chapter. Their utility is evaluated and some automation plans regarding focusing are concocted.

6.1 Focus and Modulation Transfer Function

As stated earlier in Chapter 3, MTF describes the decrease in modulation depth (aka contrast) as the spatial frequency increases (i.e. system images finer details). Since the overall MTF curve can be hard to represent in a prompt way (it is in the end a long sequence of measured numbers), an alternative way to describe MTF is required. The answer to this is MTF50, which is defined as the value of the spatial frequency where contrast has decreased to 50%. With the same logic MTF20 is the spatial frequency where modulation depth has decreased to 20%. Now the imaging properties are easily described by a single value.

It is not feasible to measure the contrast for several different black and white line strip patterns separately, since that would take long, intervals between measurement points would be too scarce and measurement accuracy would suffer. Instead the target is a gray slanted edge square (seen in Figure 6.1, which will provide MTF values for all required frequencies (remember that a square pattern consists of several frequencies as described in Chapter 3). This algorithm is used in the ISO-12233 standard [8] of measuring MTF and the steps will be explained and are summarized in Table 6.1.

The first step to measure MTF is to select the region of interest (ROI); the slanted edge is to be investigated. Then the ROI data values ($n \times m$ matrix determined by ROI) shall be linearized by inverting the opto-electronic conversion function (OECF) of the camera. This is done with methods specified in ISO 14524 [43], unless the truthfulness of the raw images captured by the camera is evinced (i.e. it is approximated that OECF outputs the input, because the sensor is optimal).

The location, slope and offset of the edge are estimated next. First, the lines of pixels perpendicular to the edge are multiplied by the Hamming window vector of the same length (m). This smooths the curvature of the vector. Then the one-dimensional (1D) derivative of the vectors are calculated by a [-1/2, +1/2] finite impulse response (FIR) filter. Finally, a linear fit to the derivative maximum values (corresponding to the edge interception point) is computed, resulting in the identification of the full edge.

In order to get the most accurate line spread function (LSF) as possible, the edge spread functions (ESFs) of each line are composed into one super-sampled ESF. This is done based on bins, whose width is (1/binning factor) $\times$ pixel, where binning factor is four. The bin where each data point is placed depends on the distance between the point and the edge, whose modulus with pixel unit does not have to be an integer (the best-fit of the edge is calculated above pixel accuracy). Amount of data values accumulated into bins are counted as well (for normalization/averaging). This process results in four times more data points along the direction of interest compared to the original image.

Binning the measurement data actually changes the effective sampling rate of the system. Sampling rate is two times the Nyquist frequency, but after the binning, both are four times bigger. Aliasing occurs for frequencies above the Nyquist frequency meaning that their effect in the signal is translated to frequencies below the Nyquist frequency [11]. This means that the results of this analysis will describe low frequencies more properly than without binning. Thus the characterization of the optics is more valid.

How aliasing really occurs with optical signals can be a bit more tricky than what it seems at first glance. Pixels have a finite area that can collect photons, which might not be the complete area of a single pixel. This is sometimes remedied with microlenses, which collect the light to the responsive area. [20] This increases the brightness of the image, but in terms of optical signal processing, information from a wider spatial object region is recorded. The point of this argument is that spatial frequencies beyond the Nyquist frequency can surface in unpredicted ways at frequencies under the Nyquist frequency. This is one motivation behind the supersampling.

The choice of the rotation angle of the slanted square is not arbitrary. It is obligatory that the angle is small enough, so that all bins will be filled with at least one data sample. If zeros appear somewhere in the signal (which would be the case if the angle was e.g. 45°), these contribute as additional signals with high frequencies [11]. One can argue that 1 px/tan(5°) $\approx$ 12 pixel rows along the edge are required, so that all bins will be filled with samples. This is a harsh and over-optimistic estimate, since the linear fit of the edge is not reliable with such a small number of rows anyway.

After binning, the point of view moves again to the edge of the square in the super-sampled data. The LSF is the 1D derivative (produced now by [-1/2, 0, +1/2] FIR filter) of the super-sampled ESF. Its size is that of the ESF and its curvature resembles a sharp peak. The LSF must be centered and its noise (caused by the physical system and the binning) needs to be disposed by applying a Hamming window.

Just like in the case of the periodic square pattern (Equation 3.2), the LSF represents a sharp detail that can be considered to consist of a huge number of different spatial frequencies. This means that computing the discrete Fourier transform (DFT) of the LSF results in the optical response as a function of spatial frequency. The general result of DFT is a vector of complex values, i.e. the OTF of the system.

MTF is the modulus of the OTF, normalized according to the zero frequency element

(since MTF is the relative change and MTF(0) = 100%). Finally, the MTF is corrected for the utilization of the discrete derivative FIR filter. This is done by a frequency-by-frequency (element-by-element) multiplication by the reciprocal of a sinc function.

Table 6.1. *The MTF measurement algorithm provided by the ISO-12233 standard [8]*

Order	**Step**	**Description**
1.	Select ROI containing the slanted edge	-
2.	Linearize image data using OECF	Raw images needed, inverting the effect of OECF might be necessary
3.	First and then final estimate of edge slope and offset	Apply Hamming window, compute 1D derivative, compute the central location of the edge for each row (column), apply linear fit
4.	Projection	Shift the image data along the edge direction to the edge of the ROI
5.	Accumulate the shifted data sampled at 1/4 of the original image sampling based on edge distance	Results in four times more data points (assuming a good linear fit of the edge and good SNR)
6.	Compute the 1D derivative of the ESF	Result is the LSF that details how sharp the edge is
7.	Shift to center the LSF vector	The peak of LSF will be at the center
8.	Apply a Hamming window to the LSF vector	Diminishes the noise and smooths the curve
9.	Compute the DFT of the windowed, binned LSF vector	Scrutinize the spatial frequency space instead of the image space
10.	Compute the modulus of the complex DFT array	MTF is the modulus of the OTF
11.	Normalize the modulus vector by the zero-frequency value (first element of array) to obtain the E-SFR	MTF(0) should equal 100%
12.	Correct the MTF for the discrete derivative response	-
13.	Report MTF result and derived metrics	-

From the fundamentals and details discussed in chapters 2 and 3 we arrive to a causality:

the object plane of the HMD Eye (with respect to the wavelength in inspection) is the plane which results in the highest MTF50 value when imaged by the HMD Eye. The most important region of imaging is the center of the image (i.e. the objects close to the optical axis). This means that the paraxial approximation is well justified to use in the measurement scenario.

The camera is joined to the optics by C-Mount threads (32 rotations per inch). The camera sensor is moved to the required image plane by rotating it along the threads. The first step of calibrating the optimal focus is to find an initial, approximate focus by rotating the camera. The camera is focused enough when the edge of the target square is as sharp as feasibly possible without.

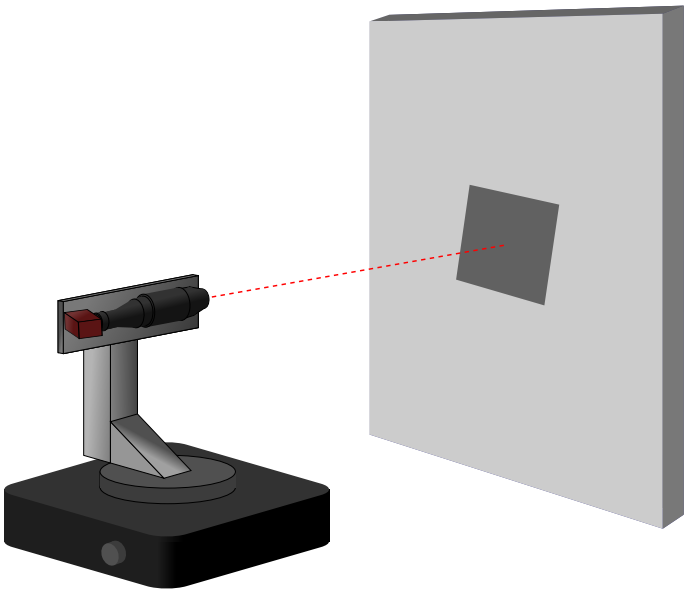
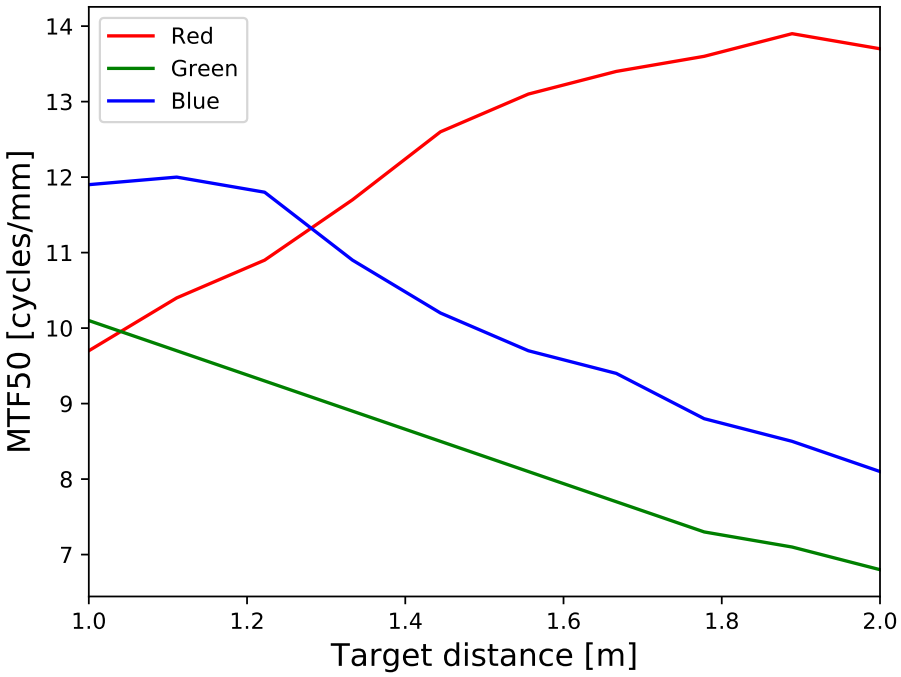


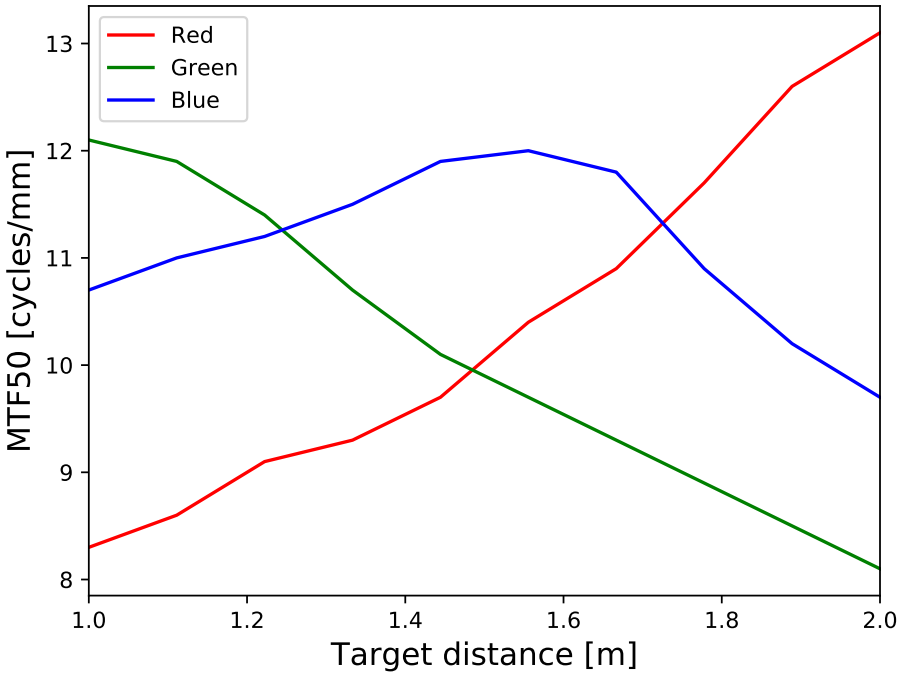
Figure 6.1. *Schematic setup for measuring MTF. HMD Eye is mounted on a motorized goniometer, which is aligned so that the optical axis intersects the center of the slanted square. The slanted edges of the square are what is measured.*

Next the MTF of the target is measured and stored (illustrated in Figure 6.1). This iteration is continued by moving the target and switching illumination wavelength and camera exposure time. The result is the MTF50 as a function of target distance and wavelength. From these plots it is possible to infer the current focus of the HMD Eye (e.g. Figure 6.2a).

Equation 2.5 can be used for focus calculation, as long as the known effective focal length of the HMD Eye is used and object and image distances are measured with respect to the principal planes. The first primal plane of the HMD Eye is very close to the aperture (around 7 mm away), so the object distance is simply the distance between the target and the HMD Eye. The image distance is changed by rotating the camera by an amount predicted by the imaging equation. The result of doing so is shown in Figure 6.2b. Figure 6.2 resembles the concept of depth of field, i.e. how much the target can be moved while maintaining good enough imaging quality.



(a) *A realistic mockup of a previous MTF50 measurement iteration, where the uniform focus was 1.1 m.*



(b) *A realistic mockup of an MTF50 measurement iteration after moving the sensor to a position corresponding the uniform focus of 1.5 m.*

Figure 6.2. *The end results of two consecutive MTF50 measurement iterations.*

The effective focal length of the HMD Eye is 7.3 mm and the object distance can be set between 1 m and infinity. This means that the camera sensor location needs to vary within a distance of around 0.05 mm (calculated from Equation 2.5), responding to a rotation area of around 24°.

When the iteration is finished and the requested focus found, the HMD Eye's camera is locked into the mechanics. Caution is a priority in this calibration, since the slack and imprecision of the C-mount thread can result in a faulty orientation. The tightness of the threads can cause stress that results in a small misalignment of the camera sensor (which can be crucial considering the 0.05 mm movement range).

Zemax [44] simulations and known features of the HMD Eye presage that the focus of the HMD Eye for different wavelengths manifest in the object space in the following order: green, blue and lastly red (see Figure 6.2). A common scenario is that the MTF50 values must be as high and homogeneous as possible with all three wavelengths. As already established in Chapter 3, this is hampered by chromatic aberration.

The MTF50 values for green light start to descent underneath the MTF50 function of blue light slightly before its maximum. The MTF50 function for red is ascending and crosses that of blue's after the maximum of blue. This is shown in Figure 6.2.

The conclusion is that the requirement of excellent and homogeneous imaging between the three different wavelengths is satisfied when the HMD Eye's focus is around that of blue light, if small distinction between the values is acceptable. This is a uniform focus, as shown in Figure 6.2. Another scenario is that only one wavelength is considered and the MTF for this wavelength must be maximized.

The MTF is very dependent on the input field. This means that in many optical systems the imaging quality drops tremendously when imaging objects departed from the optical axis even in the scale of few degrees. Since the MTF target (grey slanted edge square) has a finite size, caution is required so that MTF values of field of interest are actually measured. The slanted edge image actually consists of inputs from an angle range, and the measured MTF is more like an average of these values.

6.2 Measuring and Correcting Geometrical Distortion

Distortion is possible to be corrected with software quite easily, since the image quality is not decreased by it in principle. The locations of the image points are dislocated, and after the correction the locations should correspond to a constant M_T throughout the image. This constant value of M_T is the value at the optical axis (and this depends on the design of the optics).

The distortion algorithm used in this study is the one utilized in OpenCV [15]. This algorithm requires a dot matrix target, i.e. a white background and a matrix consisting of black dots spread evenly. Dots must be separated from background by a black contour.

The dot matrix is then dynamically thresholded (i.e. nonuniform illumination is taken into account). From this binary data (dot area equaling zero and background equaling one) the dots, being distorted to ellipses, can be detected. After the dots are detected, their distortion can be measured by fitting a distortion model (center point and distortion coefficients), described by Equation 3.6. Several images (from different angles) are analyzed and the model that results in the best overall fit is considered as the correct distortion.

In the first analysis, the parameters that can be tuned in the optimization engine are the center point and radial and tangential coefficients. This is because the first step is to assure that there is practically no tangential distortion in the HMD Eye. After this the model with only radial distortion is fitted.

Modeling the distortion of the HMD Eye is important, since the overall distortion is the combination of those of the HMD Eye and the DUT in investigation. The measured target in the HMD IQ is a quite similar dot target (that is the output of the DUT screen) as that of the calibration measurement. In the DUT measurements first the captured image is distortion corrected before fitting the distortion model for the DUT.

6.3 Dark Frame and Flatfield Correction

The images captured by the system can be corrected with the flatfield and dark frame of the system. The corrected image pixels are calculated with Equation 6.1 [19]:

$$C_{\mathrm{i}}(u,v) = \frac{[O_{\mathrm{i}}(u,v) - D(u,v)]F_{\mathrm{max}}}{F(u,v) - D(u,v)}, \tag{6.1}$$

where C_{i} is the corrected pixel value, O_{i} is the original pixel value, D is the dark frame value, F is the flatfield value, F_{max} is the maximum value of the flatfield matrix/image and (u,v) is the (image) pixel location. All the aforementioned values are an average of a set of given images.

Dark frame is relatively easy to measure. All that needs to be done is that all incoming light has to be externally blocked while capturing a set of pictures with the camera. Flatfield measurements require a plane target with a uniform angular intensity distribution profile that is placed just before the lens so that it is out of focus (as described in Chapter 3). There exist many such light sources, but in this study Image Engineering CAL2 Illumination Device [45] was chosen. When the optical axis of the HMD Eye is perpendicular to the source plane and the HMD Eye is as close as possible to the source, a set of images is captured.

The measurements are not a big burden, but it is important to understand that the dark frame and flatfield are also functions of temperature and exposure time. The measurements need to be done so that they correspond to the real use case.

It is valuable to keep in mind that flatfield correction does not hinder the effects of infor-

mation loss caused by possible stray light. If the optical system is susceptible to stray light created by rays outside FOV, the flatfield image should be taken of a distribution that is limited to FOV. This is suitable if the system will be operated within the fields limited by FOV (as is the case with almost all DUTs with the HMD Eye).

6.4 Exposure and Power Linearity

The linearity of the illumination properties, both exposure time and input power, are measured with an integrating sphere photometer (ISP). Integrating sphere is a hollow sphere with a diffusely reflecting internal surface with a couple of entrances for the light sources, optical detector and output [46]. These together form the ISP.

The output of the ISP depends only on the input power (i.e. no remarkable losses). Very small losses happen with light that traverses back to the sources or to the sensor. The radius of the integrating sphere is designed to allow quite high FOVs.

The FOV of the HMD Eye is large enough that the optical detector (in the used ISP) is visible in the images captured with the HMD Eye. This does not deter the process, since its effects in the average illumination of the image ROI is small.

In the exposure linearity measurement, the input power of the ISP is constant. The exposure time is varied from very small values to higher values, until saturation occurs (captured pixel intensity cannot grow higher). For each image the average illumination is measured and plotted as the function of the exposure time. A linear fit is modeled (with least squares approach) with its trend and highest deviation.

The power linearity measurement is executed quite similarly, but the variable in the measurements is the input power of the ISP. Linear regression is also used in the modeling phase.

The spectral radiance values are converted to the photometric luminance values that correspond to how humans perceive spectral radiance. Luminance values are that of interest in the HMD IQ. The radiance-to-luminance conversion equation is of the following:

$$L_{\mathrm{v}} = K_{\mathrm{m}} \int_{360\ \mathrm{nm}}^{830\ \mathrm{nm}} L_{\mathrm{e}}(\lambda) V(\lambda) d\lambda, \tag{6.2}$$

where L_{e} is the radiance, $K_{\mathrm{m}} = 683$ lm/W is a scaling factor between radiometric and photometric quantities and $V(\lambda)$ a Commission Internationale de l'Eclairage (CIE, translated as International Commission on Illumination) standardized spectral luminous efficiency function for photopic vision [47]. The luminance measurements are done only for a narrow green light source, meaning that $V = 1$ and thus the defined integral is equal to L_{e}(535 nm).

The measured exposure and input power ranges should correspond to the DUT output

(i.e. it is futile to measure output powers the HMDs cannot and shouldn't provide). Understanding the effects of the camera's OECF to the raw images is required. The cameras used in HMD Eye are of great quality (and tested by the manufacturer), meaning that the OECF and linearity properties are not a cause of trouble.

6.5 Optical Axis and Rotation Symmetry

The optical axis calibration is chosen to be based on the distortion center and flatfield center methods described in Chapter 3. The main reasons for this are that the recognition of the distortion center is carried on in the distortion measurements of the HMD IQ, and the flatfield gives crucial information about the interchangeable aperture at the tip of HMD Eye.

The flatfield centers are computed by thresholding and applying a Gaussian blur to the flatfield, then fitting a circle whose diameter is less than half of the shortest side of the image but more than one sixth of it. The center of curvature of this circle serves as another measure for the optical axis. However, it is not used in the analysis process of HMD IQ, but only in the verification phase.

Rotation symmetry is measured by the comparison of the standard deviation of the radial distortion coefficients. If the variance of M_{T} exceeds a given amount, rotation symmetry might be questionable.

6.6 Calibration Automation

The process of measuring the MTF as a function of target distance and wavelength is quite arduous and has the human factor in terms of accuracy. Motivated by this, the following automation system is proffered. The system consists of a virtual image forming lens, a linear motor that changes the virtual image distance and automatic illumination and image capture (with finetuning).

With the automation system, virtual image of the MTF slanted square edge target is formed to different distances between one meter and infinity (this range is defined in the HMD Eye specifications) by moving the real object. This is shown in Figure 6.3. The virtual image distance is calculated with the imaging equation, where the wavelength dependence of the virtual image forming lens focal length is taken into account. The virtual image of the target serves as a real object for the HMD Eye.

The measurement is otherwise the same as described in Section 6.1, but now the target is the virtual image of the physical target. From Equation 2.5 it is possible to derive that the image is virtual for a lens with a positive focal length when the object distance is less or at most the focal length. The focal length of the system should thus be concordant with the dimensions of the linear motor used for moving the target.

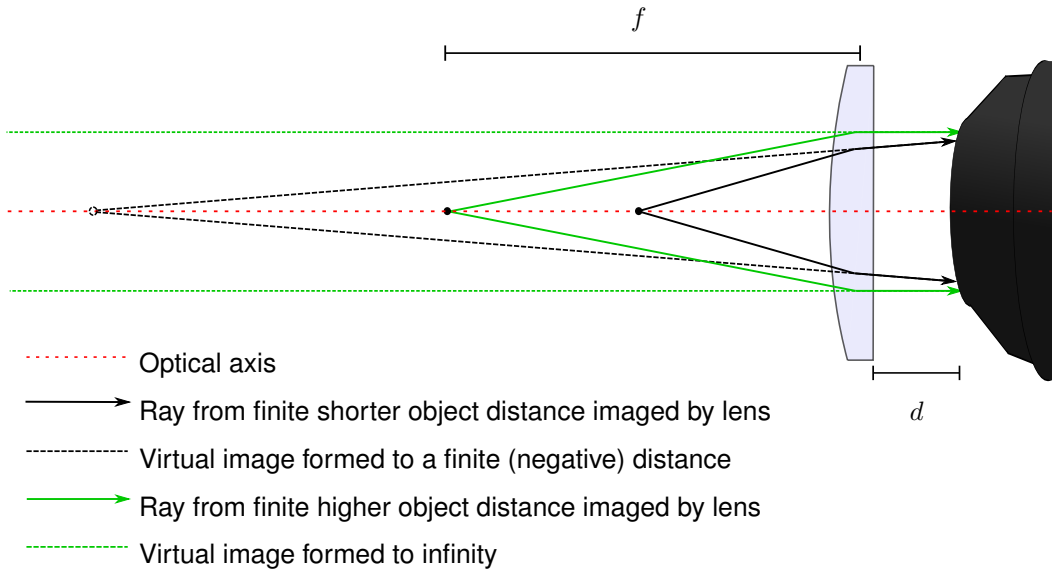


Figure 6.3. *Conceptual optical setup for providing an exchangeable virtual image, seen as a new object by the camera. f is the focal length of the virtual image forming lens and d is the distance between HMD Eye and that lens.*

When the first iteration including all its steps is complete, i.e. MTF50 is measured in all the distances with all wavelengths, the current focus of the HMD Eye can be inferred. The distance between the aperture and the virtual image forming lens is very small. Since the virtual image distances will vary between 1 m and infinity (according to the HMD Eye specifications), the object distance can be approximated as the virtual image distance.

In order to alleviate the trouble of understanding the complete operation of the system, two diagrams created with recognized tools, UML [48] and C4 [49], are introduced in figures 6.4 and 6.5.

This system provides also other functionalities as well. The opening angle the target is seen at changes less when a virtual image is used. This is because the longitudinal and transverse magnification are dependent on each other: when the magnitude of the virtual image distance grows (seen as more negative i.e. further in HMD Eye's object space), so does $|M_\mathsf{T}|$ (as seen from Equation 2.8). Here it was assumed that $M_\mathsf{L} \approx \Delta x_\mathsf{i}/\Delta x_\mathsf{o}$.

The FOV of a virtual target is thus:

$$\mathsf{FOV} = 2\arctan\left(\frac{|y_\mathsf{i}|}{|s_\mathsf{i}|}\right) = 2\arctan\left(\frac{M_\mathsf{T}h}{-s_\mathsf{i}}\right) = 2\arctan\left(h\frac{f/s_\mathsf{o}-1}{f-s_\mathsf{o}}\right), \tag{6.3}$$

where h is the greatest dimension of the target.

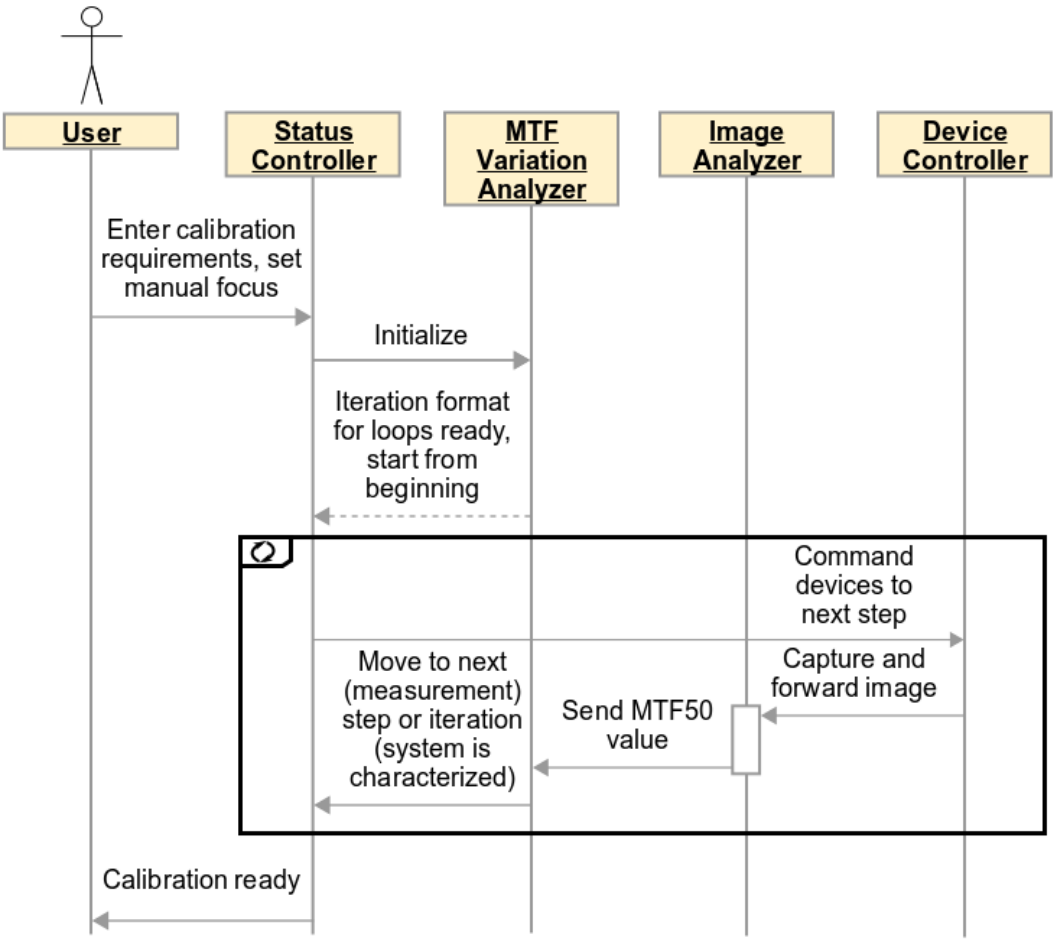


Figure 6.4. *UML [48] sequence diagram of the MTF calibration automation system*

For virtual image forming lens having f = 400 mm and target dimension being h=10 cm, this corresponds to a FOV range between 38.5° (at 1 m) and 28° (at infinity). In real object space a target of 79.5 cm corresponds to the FOV of 38.5° at 1 m, while at 15 m ($\approx \infty$) this corresponds to 3°. This has a high influence to the MTF values, possibly leading to false conclusions.

The automated calibration system would greatly benefit from motorized sensor movement. Not only would it be possible to calibrate the system without user intervention, but the possibility of measuring the focus (at given object fields) with many sensor positions corresponding to the region of 1m and infinite object distances can be a huge improvement to the new possibilities related to the characterization of DUTs (physically and software-wise). A pragmatic example of this is measurement of the object distance of the NED virtual image and the three dimensional warp of the respective object plane.

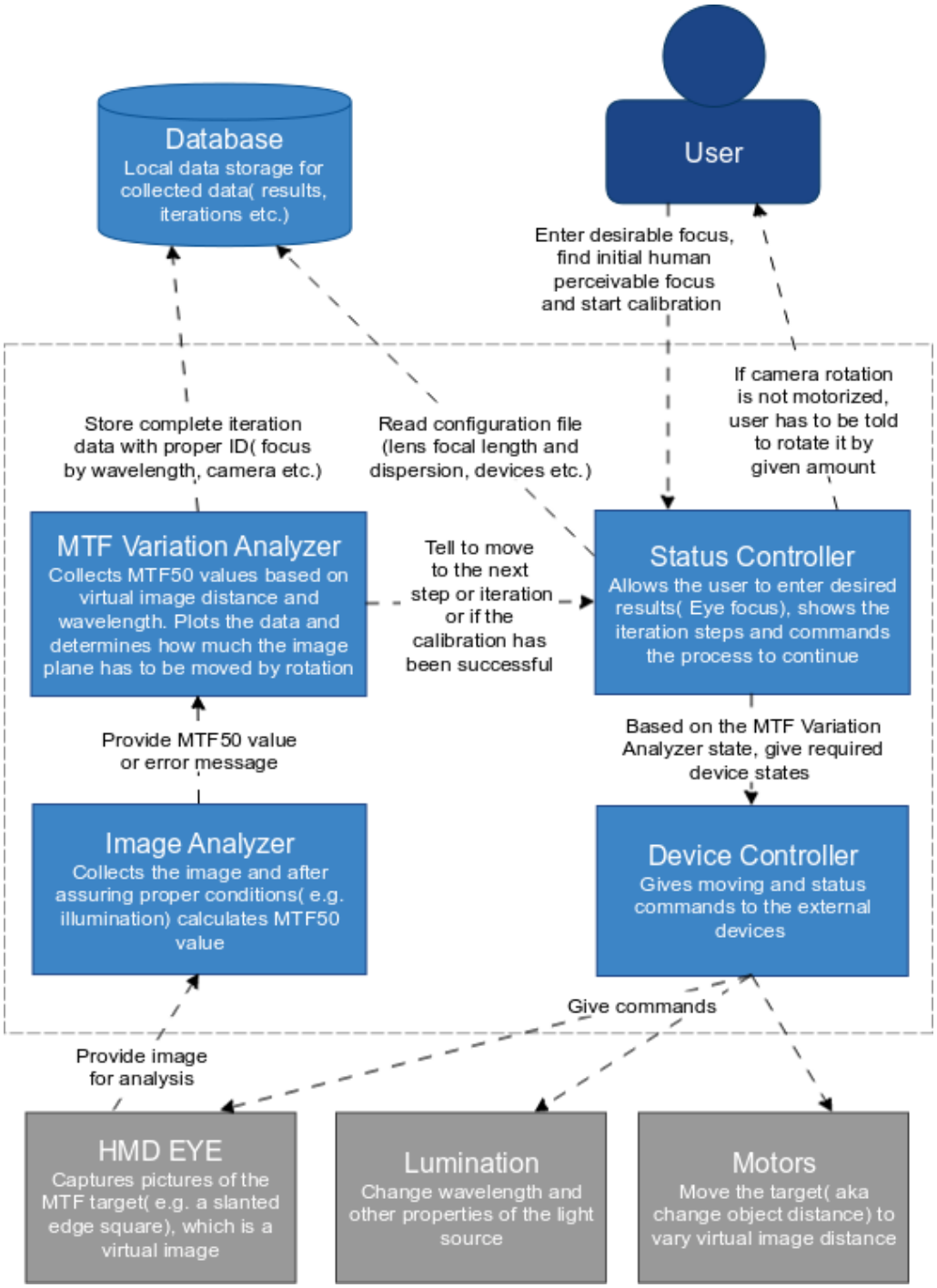


Figure 6.5. *C4 model [49] container diagram of the MTF calibration automation system.*

Causion is required when using the slanted edge algorithm at higher object fields. Since distortion can results in edges that are not linear, the basis of supersampling becomes questionable. Effects of distortion correction to the MTF calculation are not known, and they are not concerned in the ISO-12233 standard. The effect of focus on distortion properties is considered minimal based on the optical architecture of the HMD Eye.

Other measurements could also be automated further, but the value of doing so is considered minimal. For this reason, no additional automation is planned.

7 EXPERIMENTAL RESULTS

In this chapter, raw and direct measurement data is shown for both the HMD Eye and a given VR headset. The measurement equipment is also examined.

7.1 Results for HMD Eye

The wavelength spectrum of the led light source (utilized in target illuminations in slanted MTF and distortion measurements) is displayed in Figure 7.1. The peak wavelengths are approximately 457 nm, 518 nm and 643 nm. As can be seen, the bandwidth for each color is quite wide. The full width half maximum (FWHM) is approximately 25 nm for blue, 41 nm for green and 15 nm for red. The light source was thermally stable.

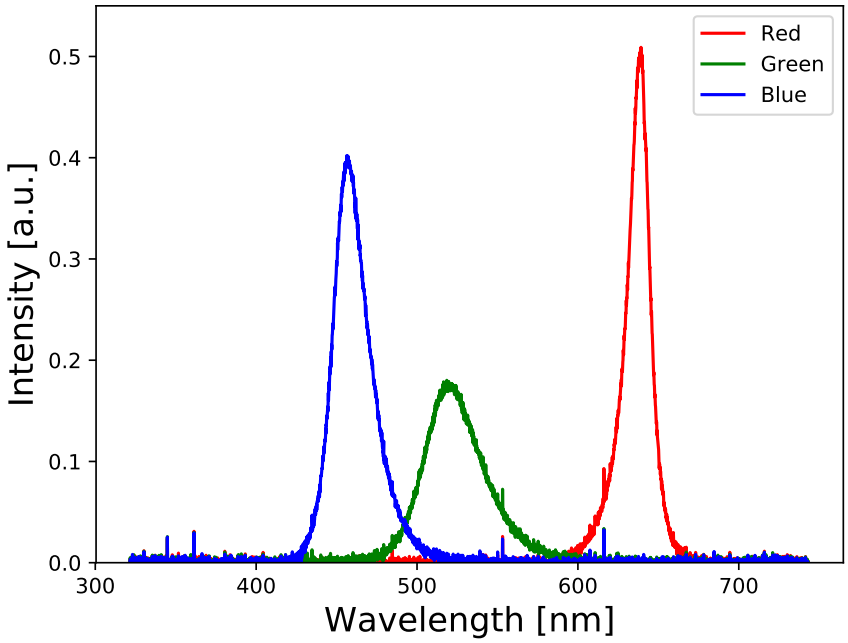
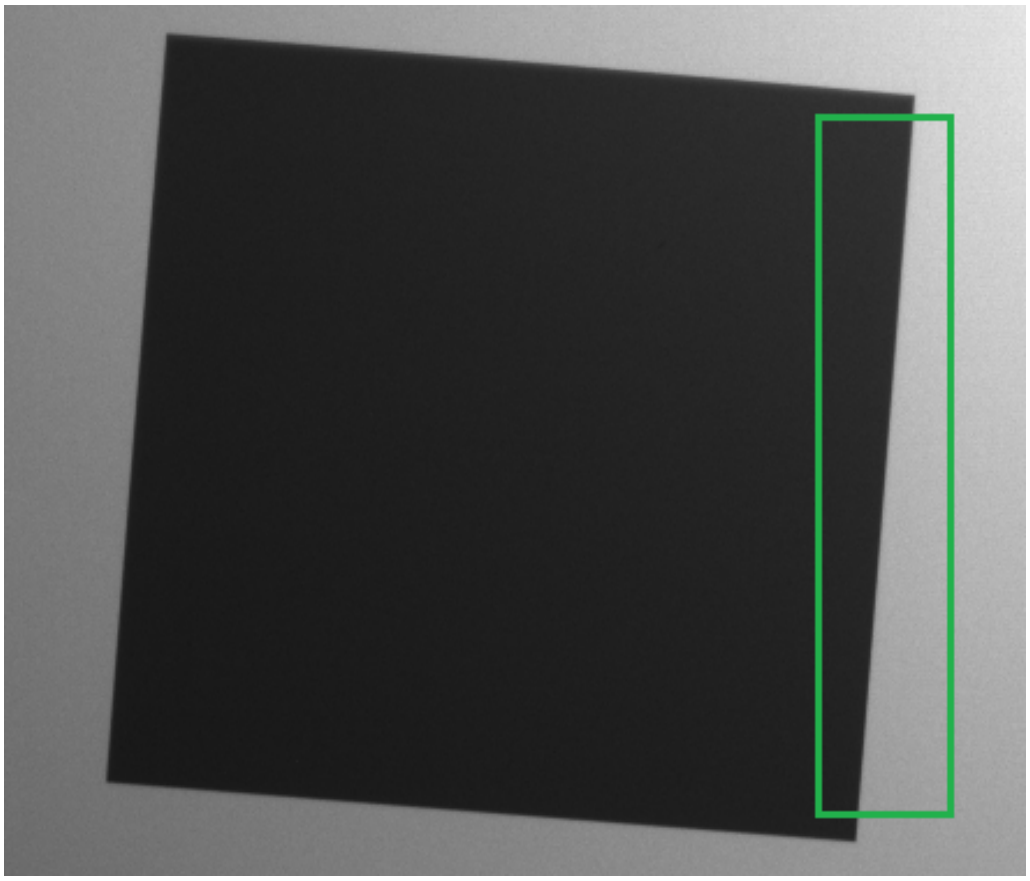
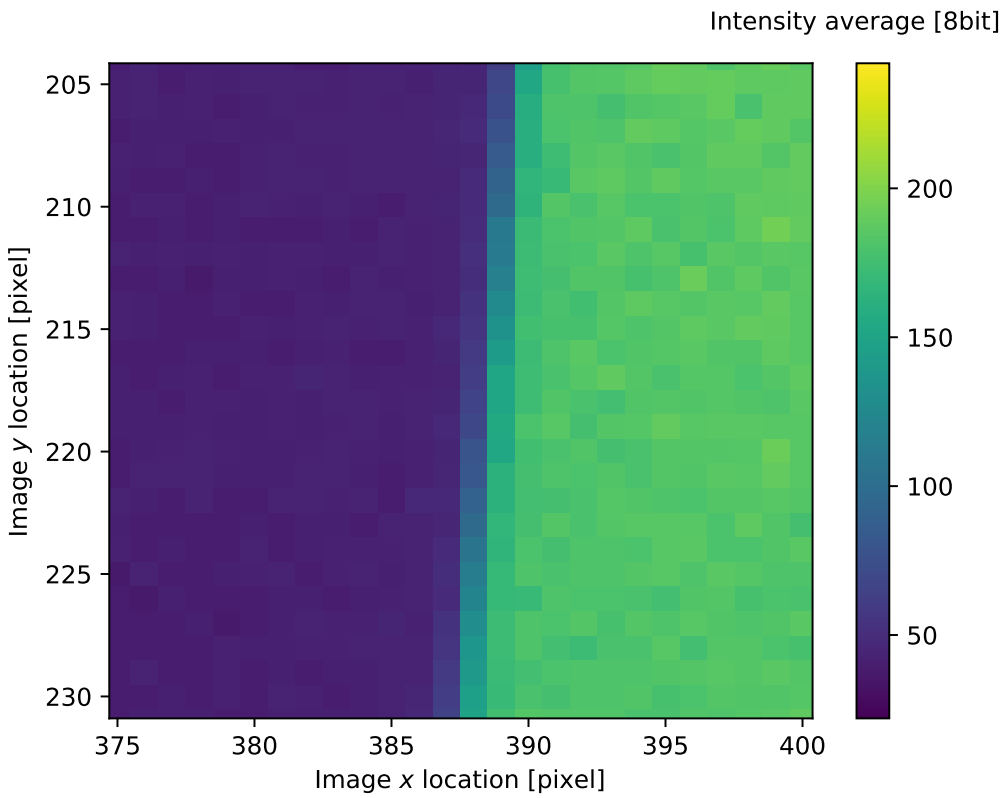


Figure 7.1. *The wavelength spectrum of the illumination light source. The peak wavelengths are approximately 457 nm, 518 nm and 643 nm.*

The random codename for the test unit HMD Eye in this thesis is 'EYE2'. EYE2 is equipped with a Manta G-1236B camera (4112x3008 sensor with pixel size of 3.45 μm) and a 3.6 mm aperture. The slanted edge square data is shown in Figure 7.2.



(a) *The complete square with a slightly nonuniform illumination. The green rectangle confines the measurement data ROI, i.e. the sample.*



(b) *Measurement data at the center of the right edge.*

Figure 7.2. *Slanted edge square target imaged by EYE2 with blue illumination that of Figure 7.1. The polar object field angle at the middle of each edge is roughly 5°.*

The slanted edge square was illuminated with the light source depicted in Figure 7.1. The data presented in this thesis is for testing different characterization methods. This means that the imperfect light source is fine for this task.

Distortion was measured with 12 images (yaw varying within 12° and pitch within 15°) for each wavelength. Illumination light spectra are that of Figure 7.1. The result is the OpenCV distortion coefficients, which are used to plot the complete trend of distortion.

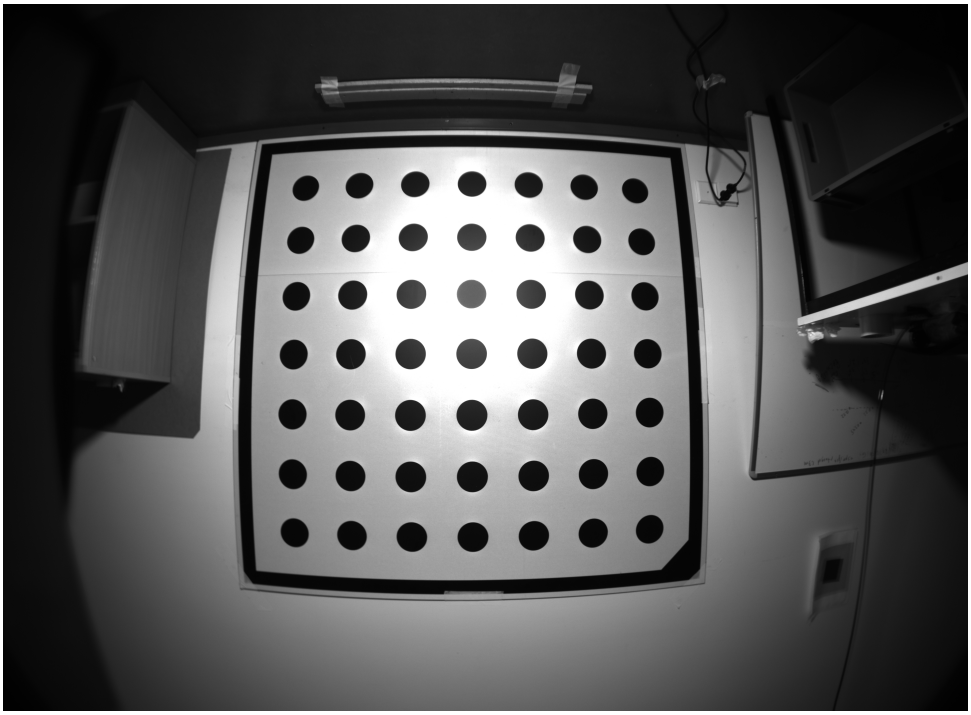


Figure 7.3. *The dot matrix target used in distortion calibration. The target was captured also in 11 other angles.*

The flatfield was measured for wavelengths of 447 nm, 535 nm, 629 nm and CIE Standard Illuminant D65 [50] white light. The light source was Image Engineering CAL2 Illumination Device [45]. The results for white light are shown in Figure 7.4.

The darkframes were measured for three logarithmically separated exposure times, 1 ms, 10 ms and 100 ms at operation temperature (in thermal balance). Ten images were captured for each exposure time, the result being their average. The central area of the darkframe with 10 ms exposure time is shown in Figure 7.5.

The linearity properties were measured with an ISP and a monochromator producing green light with wavelength of 535 nm . In order to measure the power/luminance linearity, it was mandatory to settle for fewer datapoints (the power is tuned with apertures and aperture diameters were limited). In exposure linearity measurements the radiance was measured in order to normalize the exposure time. The results are shown in Figure 7.6.

The noise levels and exposure and power linearity are sensor related properties. The optics do not affect these in practice (especially with low amplitude optical signals).

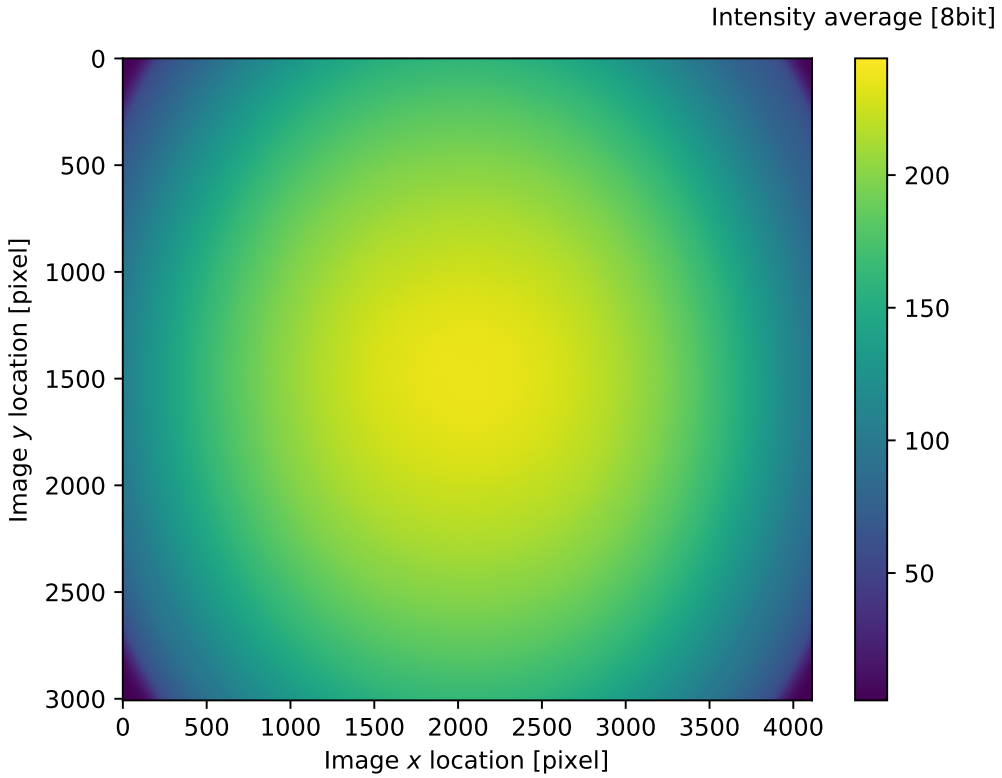


Figure 7.4. *The flatfield of EYE2.*

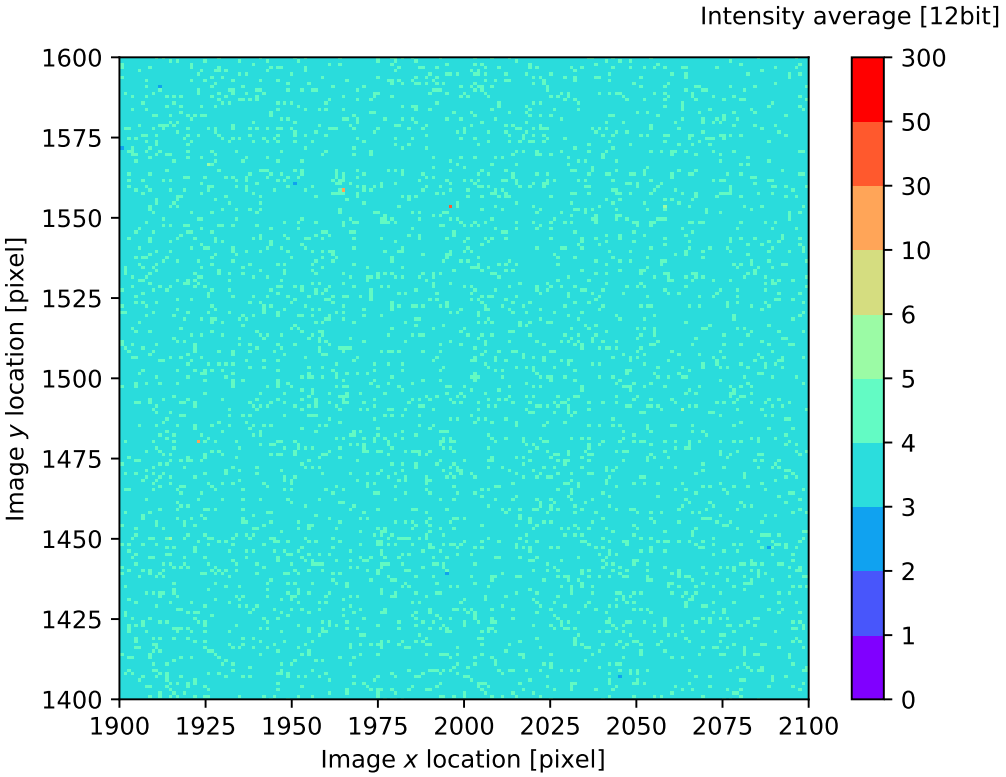
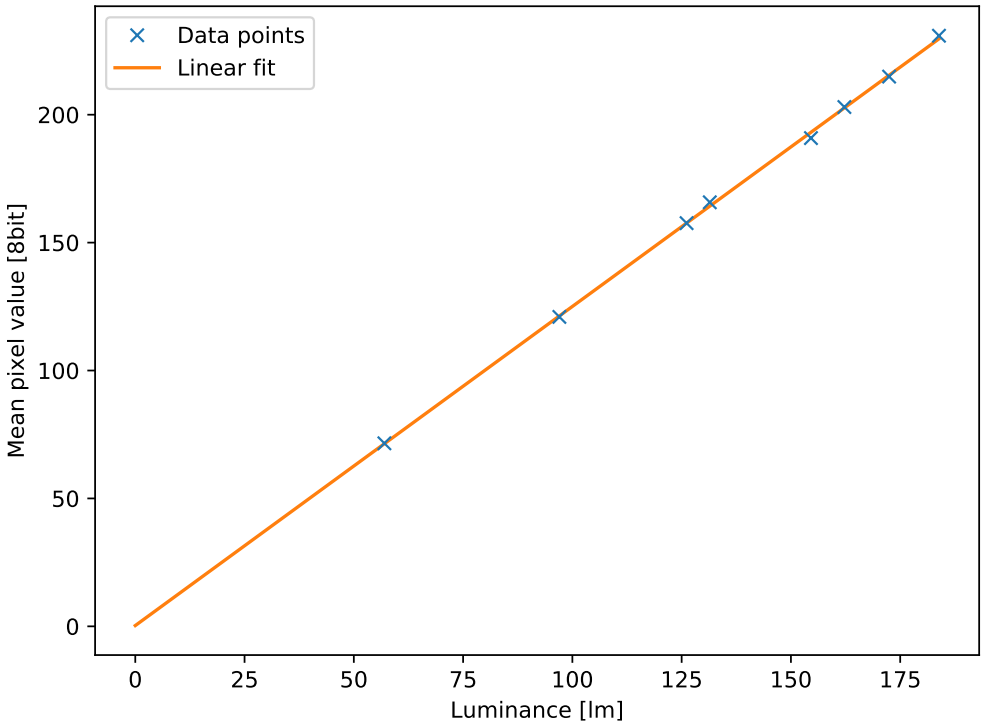
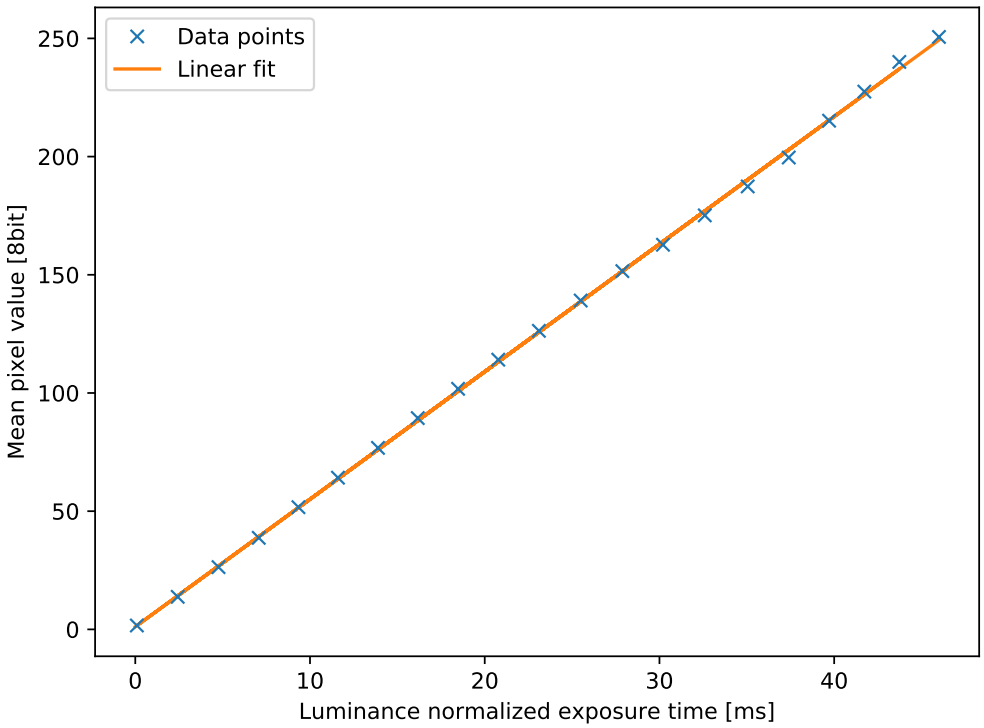


Figure 7.5. *Darkframe (at center, average of ten images) at 10 ms exposure time.*



***(a)** Power linearity of the HMD Eye (linear model $y = 1.25x + 0.30$)*



***(b)** Exposure linearity of the HMD Eye (linear model $y = 5.40x + 1.10$)*

***Figure 7.6.** Exposure and luminance (power) linearities of the HMD Eye. Data points are sparser in luminosity measurement due to the technology that provides accurate radiance output. Radiance varies a bit in operation, so in exposure analysis the exposure time has to be normalized.*

Noise levels in the darkframe and linearity properties of the sensor are measured by the camera manufacturer. These are still important to be verified.

7.2 Results for Virtual Reality Headset

The virtual reality headset used as the DUT in this thesis is a Oculus Go. Oculus Go is a stand-alone (i.e. no computer or smartphone required), budget ($\approx$ 250 €) model VR device from Facebook [51]. Pictures of this system is presentented in Figure 7.7. All the measurement data is provided by the HMD IQ system. Both left and right eye sides of the DUT were measured.

(a) *Front of Oculus Go* ***(b)*** *User side of Oculus Go*

Figure 7.7. *Pictures of Oculus Go, provided by [51] and [52] respectively*

Some specifications of Oculus Go [26] are disclosed in Table 7.1. The values in Table 7.1 should be considered critically.

Table 7.1. *Oculus Go Specifications [26]*

Parameter	**Oculus Go**
Release date [D.M.Y]	1.5.2018
Screen technology	LCD
Refresh rate [Hz]	72 or 60
Aspect ratio	8:9
Screen size per eye [pixels]	1280x1440
Approximate field of View [deg]	100
Horizontal angular resolution [ppd]	12.8

The refresh rate of Oculus Go is 72 Hz at highest and the exposure time used with the HMD Eye in the HMD IQ characterizations is higher than 10 ms. Thus, the effect of screen updates should be minimal.

First, the eyebox and FOV of the DUT were measured. This is done by turning the screen of the DUT on, displaying a completely white image. The value of all three wavelength channels in each pixel equals 255 in a completely white 8bit RGB image.

The position of the HMD eye is iterated until the effective eye relief and the dimensions of the eyebox at effective eye relief can be calculated (these two terms are refered as just eye relief and exit pupil henceforth). This is done by measuring the brightness drop in the image, intensity being a sum of the pixel values that pass a given threshold value, e.g. 95% (respect to the maximum value). Finally, a polynomial fit to the intensity as a function of aperture position in three dimensions: along optical, horizontal and vertical axes (each one being normal to one another). The output of the DUT at the center of the eyebox is showcased in Figure 7.8.

The pattern at the center of Figure 7.8 is called Moire effect and results from the edges of repetitive image features [8], in this case the spatial transitions between pixels of the DUT. The fundamental reason why this occurs is due to insufficient sampling [8], i.e. too big pixels of the HMD Eye sensor cause aliasing.

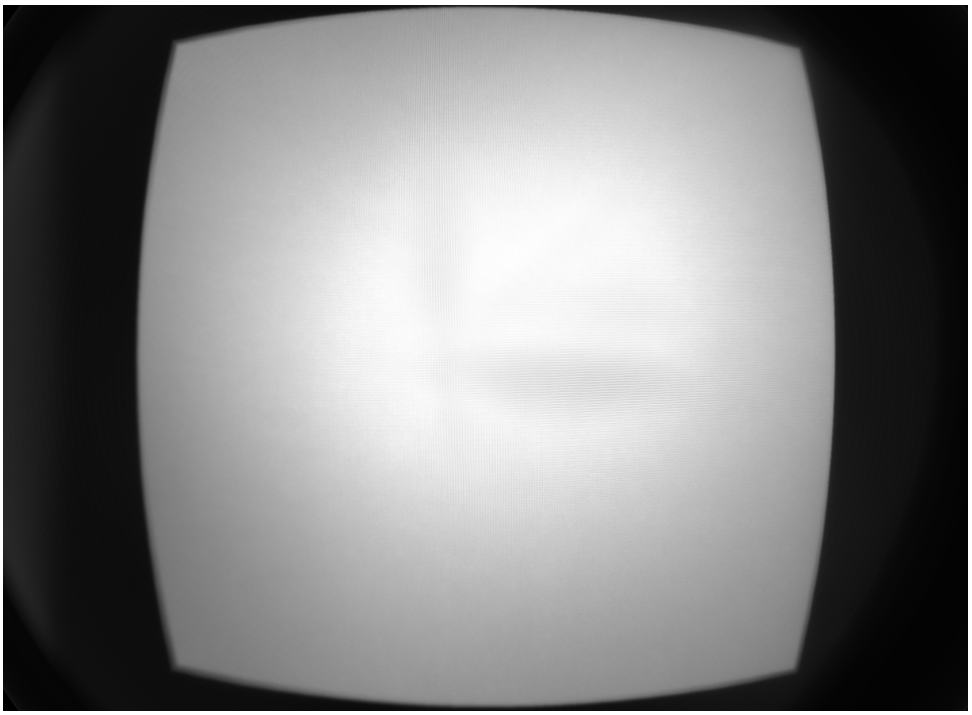
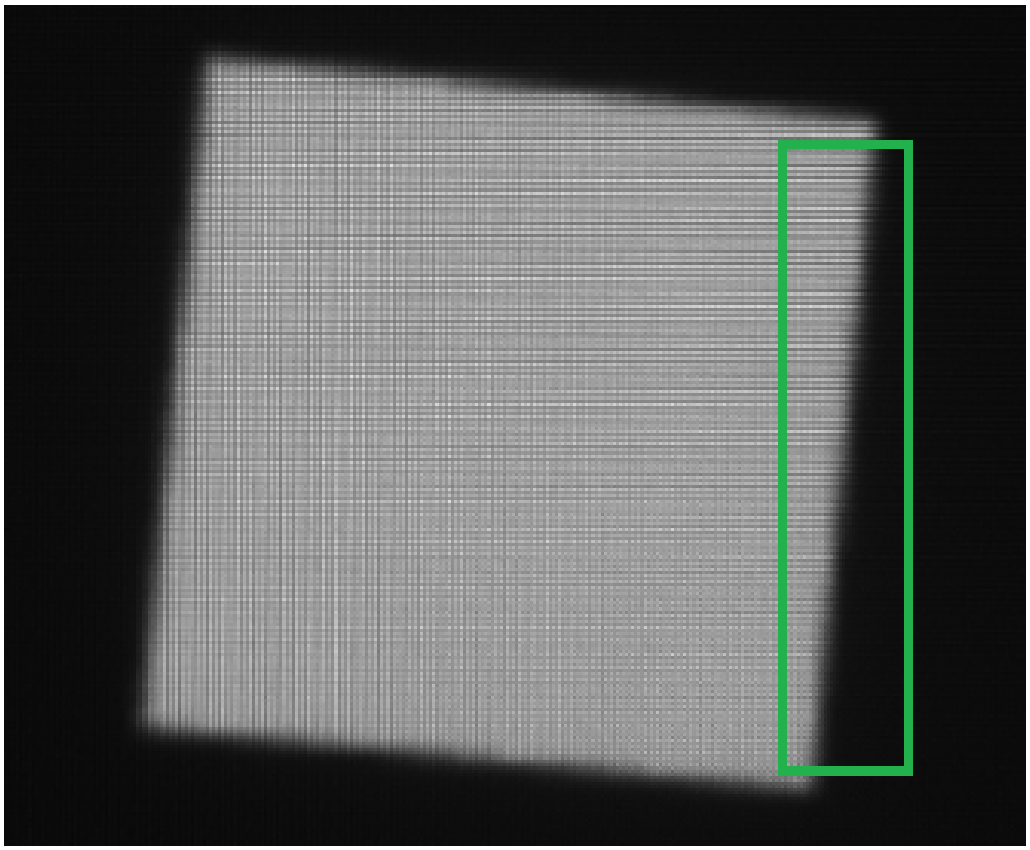
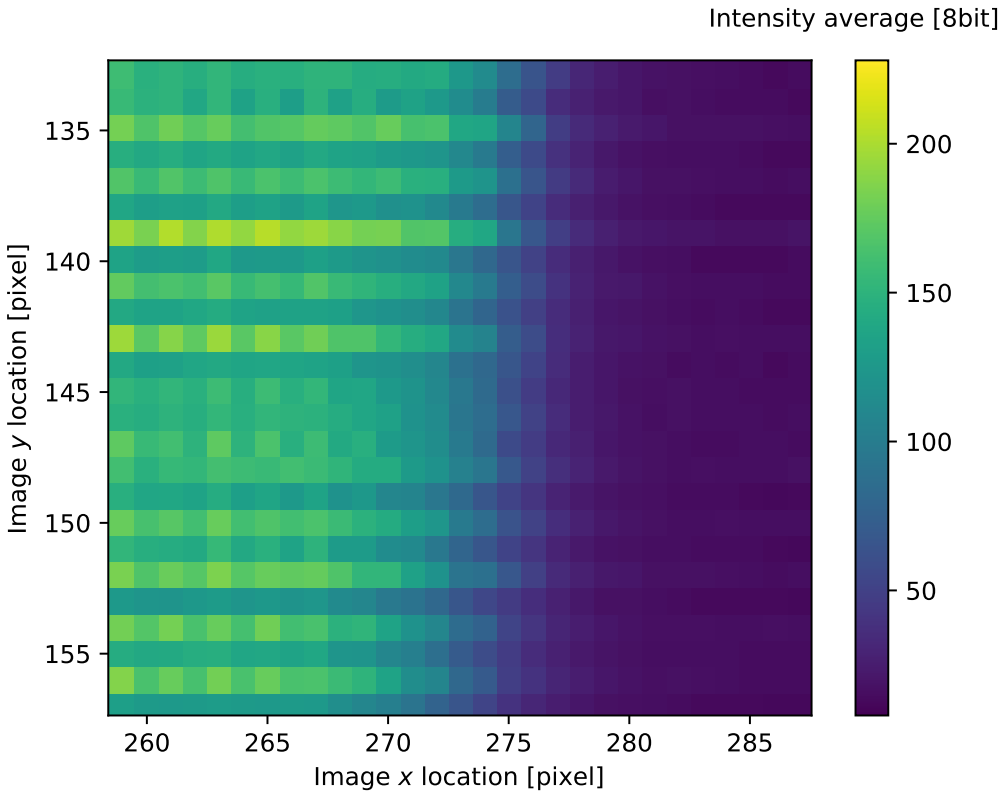


Figure 7.8. *A sample image of the white (RGB) screen imaged by EYE2 from the Oculus Go left lens. No distortion or Moire correction has been applied yet.*

After the eyebox measurements, all other optical properties can be characterized at the center of the eyebox. Optical properties are not considered at any other locations inside the eyebox. The slanted edge square measurements are shown in Figure 7.9.



(a) *The complete square from the middle of the screen, the green quadrangle confines the ROI.*



(b) *Measurement data at the center of the right edge.*

Figure 7.9. *A white (RGB) slanted edge square target imaged by EYE2 from Oculus Go left lens. The polar object field angle at the middle of each edge is between 2.5° and 5°. The Moire effect can be observed when comparing pixel rows and columns.*

The measurement data for the MTF analysis of the DUT is collected in the same manner as with the HMD Eye. The only exception is that the target object (slanted edge square) is provided by the DUT itself. Also, the contrast of the object is swapped, i.e. the square is now bright and the background is dark (one can argue that this diminishes some noise or stray light compared to most of the screen being lit). The square measurement data can be seen in Figure 7.9. The Moire effect has been taken into account in Figure 7.9.

The size of the square target in Figure 7.9a is such that the center of the edges is located approximately within 2.5° and 5° polar angles respect to the HMD Eye's angular space. Some remnants of the Moire effect can be seen in Figure 7.9b.

The white (RGB) dot grid (19x19 dot matrix, 361 dots in total) measurements are shown in Figure 7.10. It is similar to the dot target in Figure 7.3, but the distortion correction of EYE2 is applied. However, since the FOV of the DUT is smaller than that of the HMD Eye, only one target image is required. The more there are dots in the target, the more accurate the measurement in terms of interpolation.

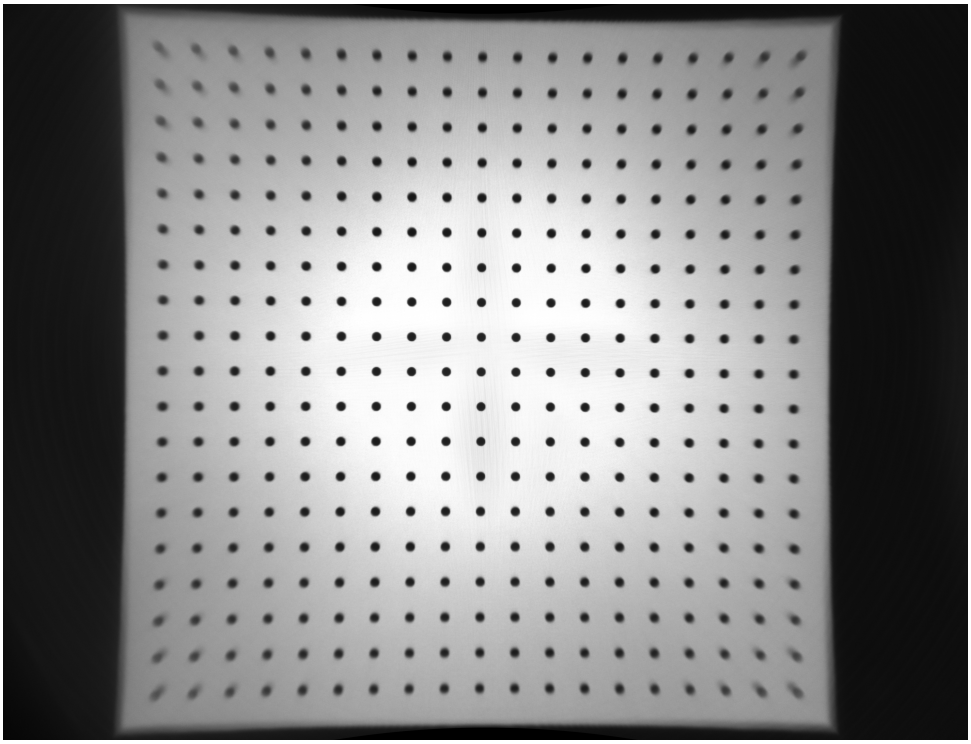


Figure 7.10. *A white (RGB) dot matrix target imaged by EYE2 from Oculus Go. EYE2's distortion correction is applied.*

The dot target is always measured so that the centermost dot is located at the optical axis of the HMD Eye. In this case, the approximation that the image center is the optical axis has been used.

8 ANALYSIS OF EXPERIMENTAL RESULTS

Lastly, the measurement results are analysed for the given test unit HMD Eye and the given VR headset. The focus of the test unit HMD Eye is shown amongst other properties, especially the MTF and geometric distortion.

8.1 Analysis of HMD Eye Measurement Results

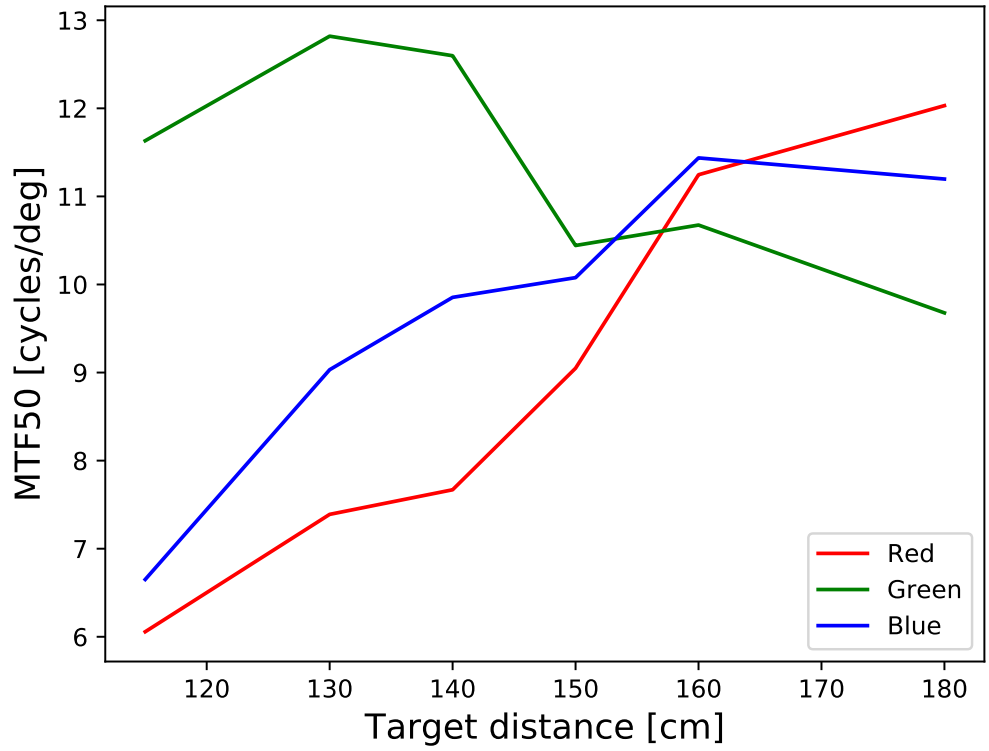
The wavelength spectra of the light source (Figure 7.1) used in MTF and distortion measurements are quite broad. This definitely has some impact to the MTF results, since chromatic aberration deteriorates the image quality, unsharpening the image of the square edge.

The initial goal was to focus EYE2 to an object distance of 150 cm. The analysis results for focus (i.e. MTF50) and MTF are shown in Figure 8.1. The original Nyquist and sampling frequencies (144.9 samples/mm and 289.9 samples/mm) correspond to the values of 18.4 cycles/deg and 36.8 cycles/deg respectively, assuming linear distortion (this is justified based on the data in Figure 8.2). Since super sampling was done, these limits no longer prohibit the analysis of higher frequencies. Frequencies up to four times the original Nyquist frequency can be interpreted as real data (with noise, of course).

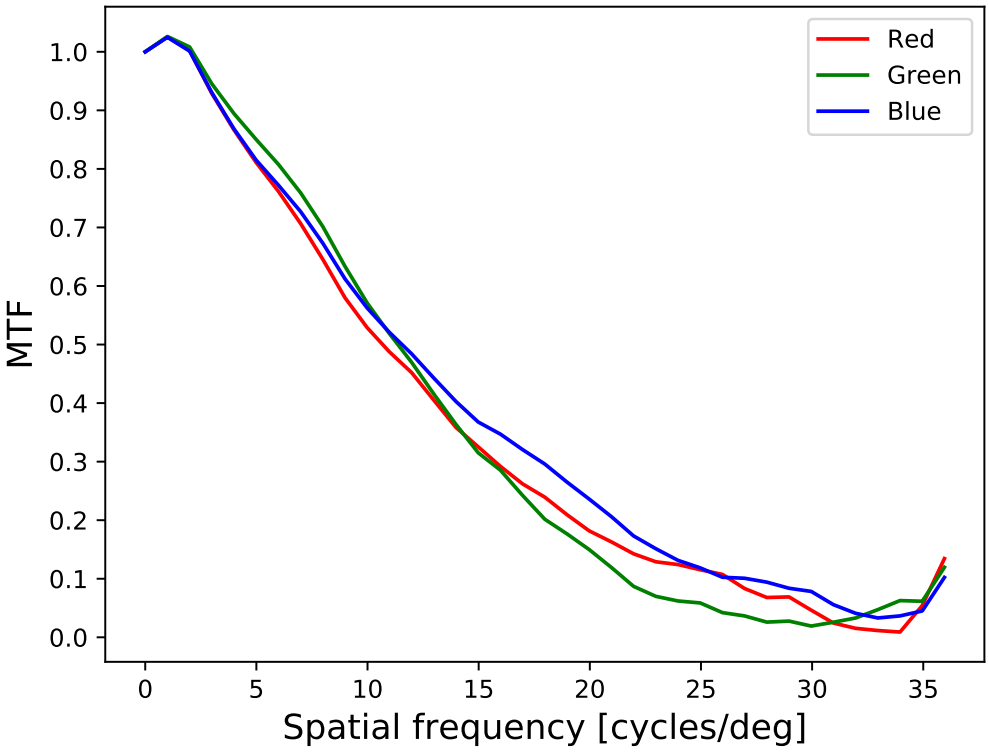
The MTF50 results (Figure 8.1a) are quite good, for starters the focus of the HMD Eye is 160 cm. Allowing a ± 10 cm error range is well justified in this case. MTF50 slowly decreases after 160 cm for blue light, whilst the value for red light continues to increase beyond the maximum value of blue.

The individual MTFs for the three wavelengths at target distance of 160 cm (closest to the real focus) are much alike and the derivative of the MTF slowly decreases. In none of these three cases does the MTF50 reach the original Nyquist frequency. If the sensor distance could be finetuned further, this limit could perhaps be reached.

Assuming that the true focus is somewhere within 150 cm and 170 cm, Equation 2.5 (with effective focal length of 7.3 mm and assuming that the first primal plane is very close to the first optical surface, which is true in this case) unfolds that the sensor has to be moved inside a ≈ 4.2 μm distance region. This is clearly unfeasible with hand rotated C-Mount threads due to the loose between the threads and lack of absolute sturdiness of the mechanics material. With motorized sensor movement this could be done.



(a) *MTF50 of EYE2 as a function of target distance and wavelength*



(b) *MTF of EYE2 at object distance (160 cm target distance) with three wavelengths*

Figure 8.1. *MTF50 and MTF measurements of EYE2*

There is a value greater than one in the MTF analysis data. One can argue that this could be due to unproper centering of the zero frequency in the curve due to the different FIR used (compared to the one used in edge linear fit detection) to calculate the derivative or due to super-sampled data causing difference. This is not true. If the zero frequency centering truly was amiss, the data at both sides from the zero frequency (negative and positive frequencies) would be equal in magnitude [11], i.e. MTF($\mathbf{k}$) = MTF(-$\mathbf{k}$). The data points next to the peak are not the same in magnitude (differing approximately by 0.005 with all three wavelengths).

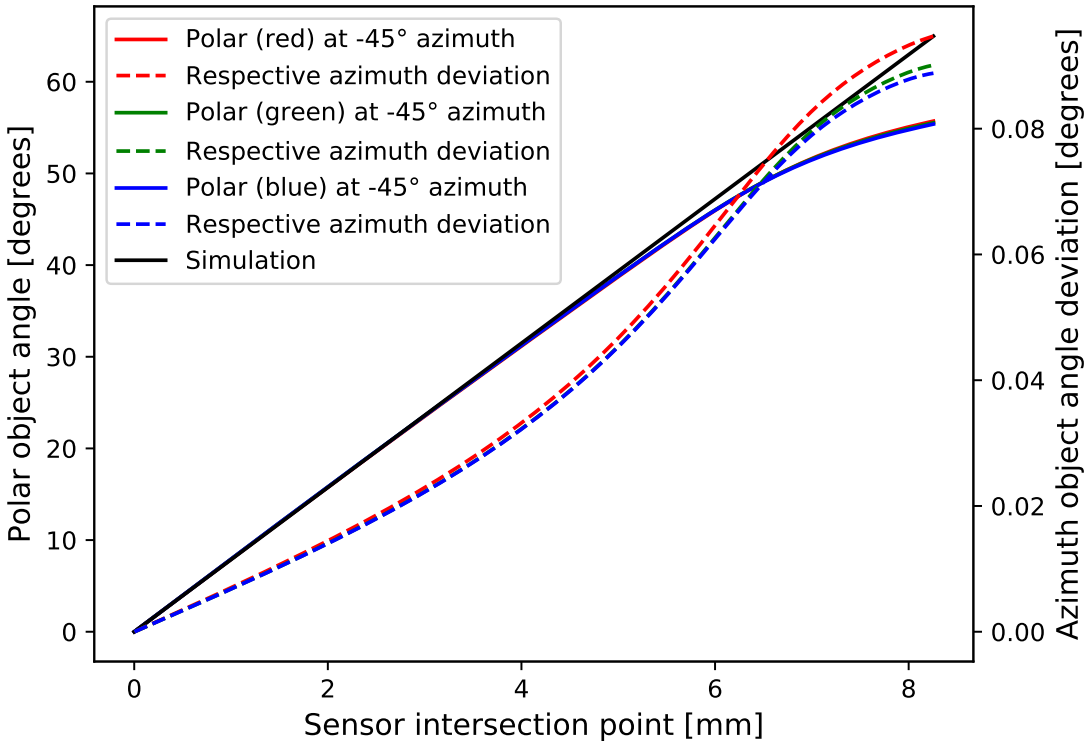
The peak can be due to static noise or simply normal behavior of the optics (some swaying is apparent in the Zemax simulations of the MTF, i.e. the derivative of the MTF can be positive at some spatial frequencies). However, the most likely case here is that the supersampling has caused slight image sharpening (not really the whole digital image, just the LSF), i.e. the contrast is much larger between pixels close to the edge compared to those away from the edge. This causes similar bumps [53] as observed in Figure 8.1b. Further investigation into the data calculation actually proves this. The source of the sharpening is most likely due to the nonuniform illumination of the square target, as seen in Figure 7.2a.

Articles [39][54][55] consider the human eye white light model for MTF based on theory and real measurements. Each model predicts that MTF50 for the human eye with pupil diameters of 2 mm and 6 mm corresponds to 15 cycles/deg and 5 cycles/deg respectively (the variance is due to diffraction and aberrations [39]). With an aperture of 3.6 mm EYE2 achieves MTF50 of around 11 cycles/deg (as seen in Figure 8.1a. This is a good result and simulates human vision quite well according the sources.

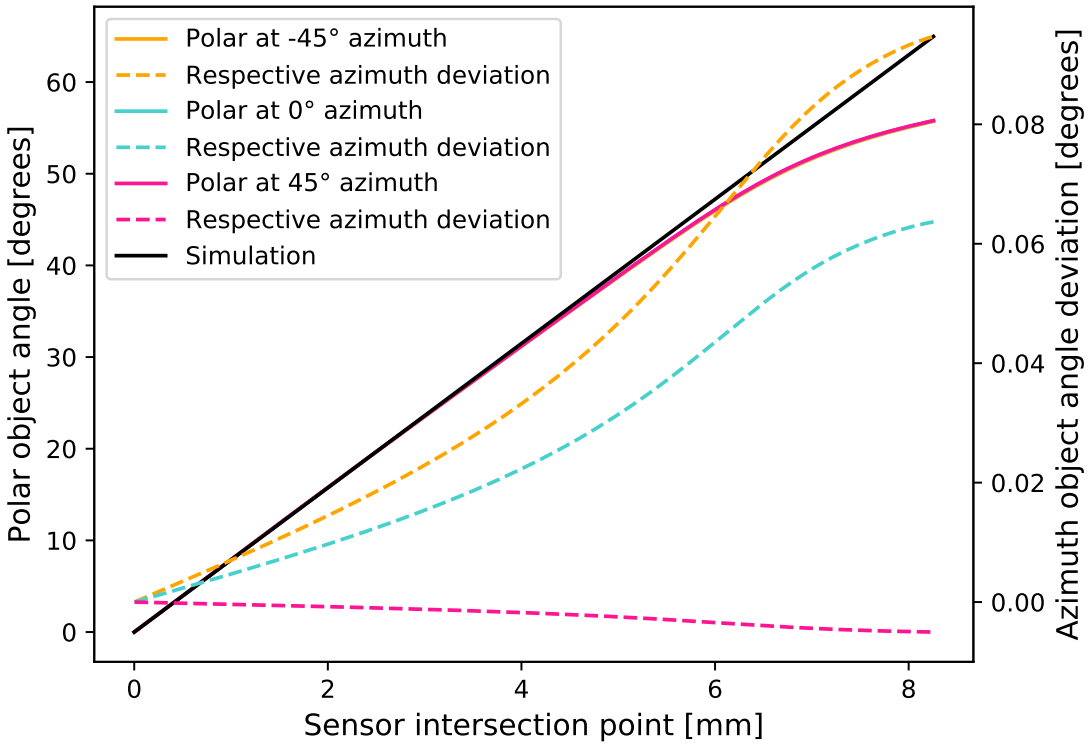
During the MTF measurements it was found out that even small misalignments (e.g. horizontal tilt) resulted in quite drastic (in comparison to how small the physical changes were) changes within the MTF values. One degree of horizontal tilt of the target resulted even in 8% of decrease of the MTF50. More sophisticated measurement setups are recommended for future endeavors with imaging property measurements.

The distortion model was calculated with k_1, k_2, k_3, p_1 and p_2 (from Equation 3.6) being nonzero, i.e. only radial and tangential distortion were considered. The reason for k_4, k_5 and k_6 being zero is that otherwise numerical instability (denominator being close to zero) can occur. Considering the results, the reliability leans to the linearity of the radial distortion, possible tangential distortion and the variance between wavelengths.

The experimental distortion data analysis is shown in Figure 8.2. Figure 8.2a shows how linear the radial distortion, i.e. polar object field, is as a function of position on the axis of (approximate) maximal tangential distortion at the sensor. The deviation of the azimuth angle in angular object space is also shown. The effect of the azimuth angle at the sensor (with blue light) is shown in Figure 8.2b. The axis of maximal tangential distortion (counter-clockwise from the positive u-axis) is -48.04° for red, -46.84° for green and -46.35° for blue (calculated with Equation 3.8).



(a) *Distortion linearity of EYE2 with three wavelengths at azimuth angle of 45°*



(b) *Distortion linearity of EYE2 with blue light at three different azimuth angles*

Figure 8.2. *Distortion linearity measurement results of EYE2. Measurement data covers polar angles up to 35°, so extrapolations exceeding 4.5 mm sensor region are invalid.*

The FOV of the HMD Eye is 129°, so only object angles up to 64.5° should be considered. However, the distortion measurement data was much more limited in respect to the angular space than what the FOV requires (the highest polar angle being $\approx$ 35°). The results show that the model is not good at extrapolating the data. The measurement data must include polar angles that reach the FOV of the system. This can be challenging, since detecting the dots and especially the contour at FOV requires a separate detection method.

Figure 8.2 shows that the HMD Eye is a barrel distorted optical system. This is self-evident for wide-angle measurement systems, since otherwise the image points would be difficult to capture on a size-limited sensor.

In the measurement region, the model is very linear (maximum relative linearity error being $\approx$ 1% close to the optical axis) and the highest azimuth deviation is 0.04° on the axis of maximal tangential distortion, meaning that these image points are dislocated towards the right side of the axis of maximal tangential distortion, approximately by one pixel unit.

The results of Figure 8.2 are in agreement with the results calculated with Equation 3.8. The small effect of tangential distortion to the polar angle can also be detected in the analysis data, if a small region around 35° polar angle is investigated. The polar angle increases in azimuth order -45°, 0° and 45°. This is concordant with the theory provided in Chapter 3, specifically in Figure 3.2c. In the very end, the effect of tangential distortion is almost diminishable.

Regarding distortion, the possible measurement and analysis errors are not considered in this thesis due to the lack of time. Errors in detection of the dots, matrix formula optimization (including the angular offset of the target matrix), possible effect of the sensor misalignment etc. have yet to be studied. The first step would be to do several recalibrations and examine the possible differences of the measured calibration models.

The flatfield of the HMD Eye (Figure 7.4) for white light (D65) is not out of the ordinary, it shows signs of vignetting like any system would. The input was angularly uniform with no limitations to the input field (i.e. light could enter outside of FOV as straylight). The flatfields change somewhat when an aperture of different diameter is introduced (this affects other properties also, so it cannot be dynamically changed without full calibration of the system). The decrease of contrast from center to the limits of FOV is around 84%, which is a considerable effect to the dynamical depth at the regions close to the limits of FOV. However, for object fields of 50° (100° Horizontal FOV) the drop of dynamical depth is 57%, meaning that approximately one bit is wasted.

The darkframe images themselves do not provide sufficient information for further analysis. For this reason, a histogram of the results over the whole sensor area is disclosed (Figure 8.3).

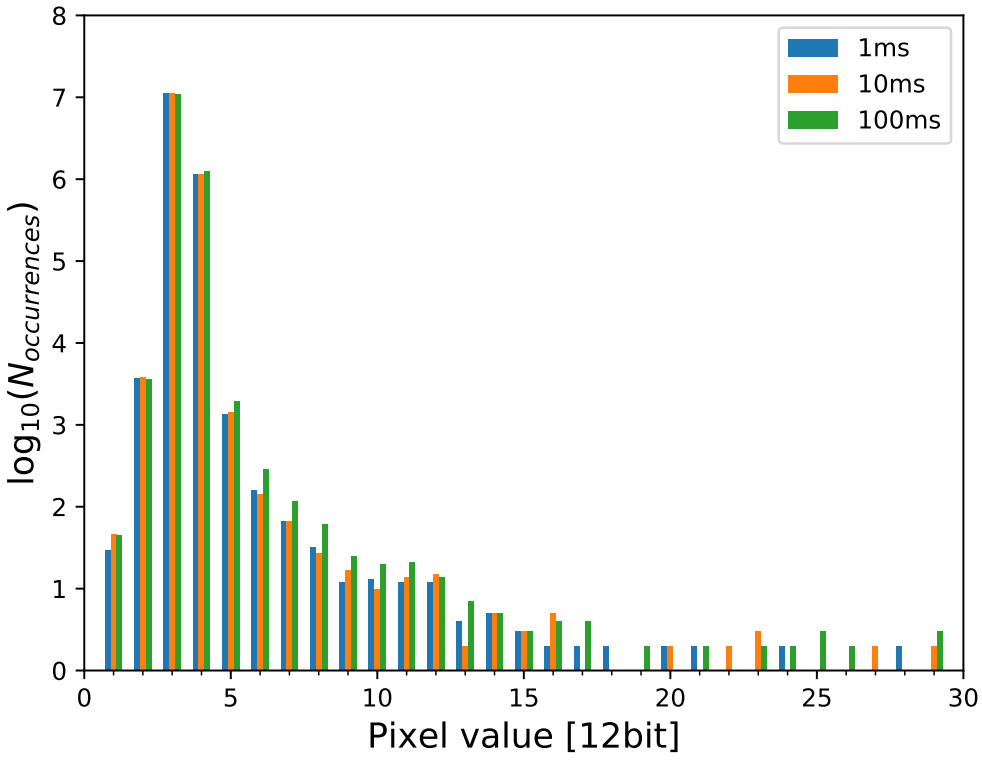


Figure 8.3. *A histogram of the intensity/dark current distribution in the darkframe throughout the sensor based on Table 8.1.*

The darkframes with each exposure time are much alike, as can be seen in Figure 8.3. They are so close to each other that arguably no further procedures are necessary when fixing dark currents with any exposure time between 1 ms and 100 ms. Even hot pixels in the darkframes (seen in Figure 7.5) are quite stable.

The mode of darkframe statistics is three in each case and its dominance does not change in practice. If a closer look to Figure 8.3 is taken the differences between the exposure times becomes apparent.

When exposure time is higher, somewhat more occurrences appear at dark current of one, but only with exposure time of 100 ms the occurences of values of 4, 5 and 6 start to increase. The latter is probably just the effect of temperature or external radiation exciting electrons whilst the anterior could be due to dark currents (i.e. the electric charge in the cells) being retired before the proper image measurement swipe.

The read noise of Manta G 1236 is 2.7 e^-/px according the specifications [37]. The temporal dark noise is reported to be 2.1 e^-/s [37]. Documentation about the level of other types of noise is vague. Since thermal noise dominates in the case of low optical signals [21] (as approximated in Chapter 3) and the entrance pupil of the HMD Eye is small, the noise at the mode and around it can be considered to be caused by thermal noise. The rest can be considered as hot pixels.

Table 8.1. *EYE2 darkframe (12bit, average of 10 images) analysis results. The table shows the occurrence* $\mathbf{N}$ *of pixel values with each exposure time.*

Pixel value	$\mathbf{N}$, **1 ms**	$\mathbf{N}$, **10 ms**	$\mathbf{N}$, **100 ms**	**Pixel value**	$\mathbf{N}$, **1 ms**	$\mathbf{N}$, **10 ms**	$\mathbf{N}$, **100 ms**
1	30	47	45	26	1	0	2
2	3769	3809	3662	27	1	2	0
3	10^7+ 1217788	10^7+ 1212041	10^7+ 1107296	28	2	1	1
4	1145587	1151234	1255328	29	1	2	3
5	1375	1432	1962	30	1	0	5
6	158	143	288	31	1	0	0
7	67	67	116	32	0	1	0
8	32	27	62	33	0	1	1
9	12	17	25	34	2	0	1
10	13	10	20	36	0	0	1
11	12	14	21	37	0	1	0
12	12	15	14	38	2	1	0
13	4	2	7	39	0	1	0
14	5	5	5	40	0	1	1
15	3	3	3	41	0	1	0
16	2	5	4	42	0	0	2
17	2	1	4	44	1	0	0
18	2	0	1	45	2	1	0
19	0	1	2	46	0	0	1
20	2	2	1	47	0	0	1
21	2	0	2	57	0	0	1
22	1	2	0	174	1	0	0
23	0	3	2	177	0	1	0
24	2	1	2	178	0	0	1
25	1	1	3				

The noise level is calculated from the data in Table 8.1 with Equation 3.9. The dynamic range at the optical axis for 1 ms, 10 ms and 100 ms exposure times is 4096 = 72.2 dB, i.e. the maximum signal divided by the minimum noise.

The dynamic range depends on the noise level and the flatfield. For this reason, the dynamic range at the diagonal FOV limit is actually 16% of that at the optical axis. This means that two bits per pixel is wasted at the very edge of the FOV, since 0.16·4096 - 3 $\approx$ 652 has to be saved as at least a 10bit integer.

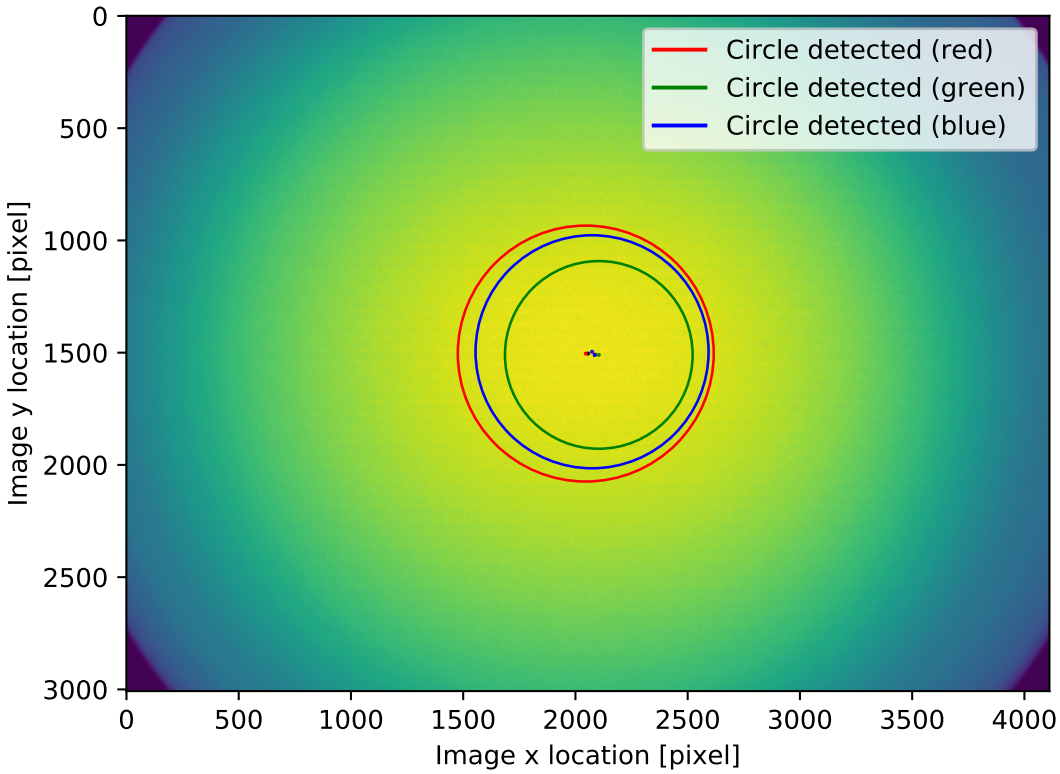
The pixels corresponding to the diagonal FOV limit can be saturated if intensive light sources are located that far from the optical axis. In the use case of the HMD Eye, this does not happen due to the quite flat lumination of the DUTs (i.e. no bright spots at the corners of the screen). Thus, in the case of the dynamic range, factoring in the effect of the flatfield is mandatory or at least promptly recommended.

Characterization of the optical axis interception point at the sensor is based on the distortion and flatfield measurements of each wavelength. Only the parameters related to the circle detection were varied between measurements for computation speed reasons (the minimum and maximum size of the diameters of the circles were constant). The distortion measurements were not locked to any result, i.e. all three wavelengths were measured independently and the optical axis was allowed to vary within computation. Optical axis interception points with the sensor are shown in Figure 8.4.

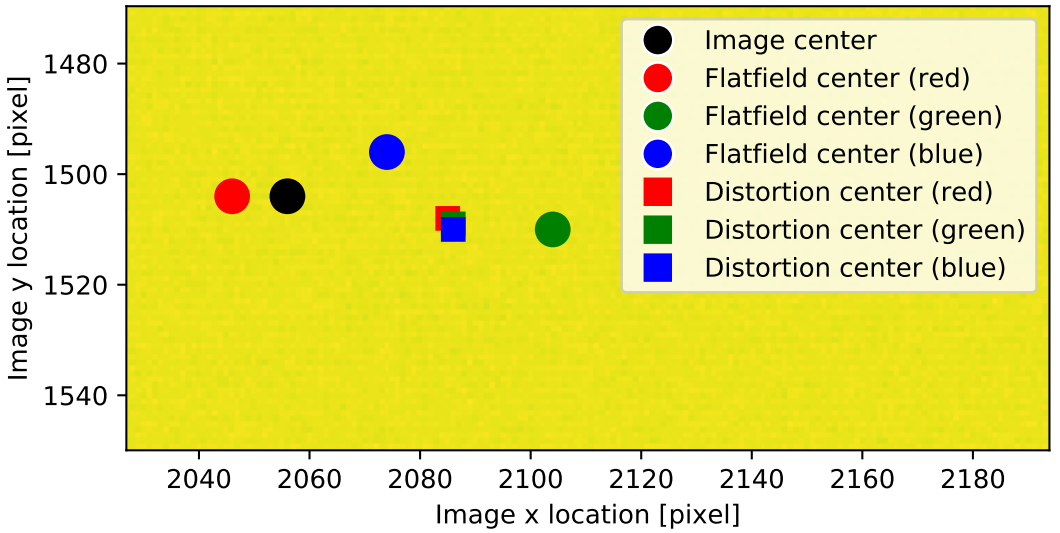
The optical axis measurements imply that the location of the interception point of the axis and the sensor is quite stable in terms distortion, difference between wavelengths being two pixel units (7 μm). Standard deviation for each measurement is 4 pixels per axis. The deviation from the center of the sensor is $\approx$ 40 pixel units. The distortion centers of the HMD Eye are really close to each other, even without using the possibility (provided by OpenCV) to lock the center between measurements. This is good evidence that uniformity reigns between the wavelengths in terms of distortion properties.

The flatfield varies quite a bit between the three cases, as seen from Figure 8.4. One fact that affects this is that the flatfields couldn't be captured with same exposure times (since the quantum efficiencies of the sensor are different, among other reasons). The circle detection is done according to the description introduced in Chapter 6.

The centers of the flatfields (aka intensity responses) vary a bit, highest deviation from each other being $\approx$60 pixel units (200 μm). This is a rather good result, considering the method of analysis. Some variance between different methods was expected. The result of the optical axis measurements is that the distortion center is the optimal reference for the interception point if such is needed.



(a) *Full view of the digital image with optical axes*



(b) *Zoomed view of the optical axes*

Figure 8.4. *The proclivity of the optical axis locations with a red flatfield image as a reference in the background. The circles correspond to the intensity distribution detection method described in chapter 6.*

The exposure and luminance linearities are really linear according to the measurement results (Figure 7.6). In both linearity measurements the maximum linearity error (the relative difference between measured and fitted data) was $\approx 0.6\%$. Good linearity was an expected result, since this has been confirmed by the camera manufacturer. It is still necessary to test that the optics do not have any influence on these properties (e.g. due to reflections, absorption etc.).

The luminance linearity measurement provides less data points compared to the exposure linearity measurement, but this is not a hindrance. The highest residual on both models is one (rounded), which is quite optimal. Both linearity models should specify that at zero luminance the pixel average should be zero in theory. This is not the case, but the models are quite close to this. When treating with very small currents in electronics distinguishing signals from noise or dark current can be an ordeal and cause uncertainty.

In conclusion, the HMD Eye under inspection in this thesis is of good quality. The focus, distortion and optical axis calibration look promising. New principles for these calibrations are laid out thanks to the results.

EYE2 still falls short in comparison to the Zemax simulation results (around 30 cycles/degree) in terms of MTF. The measured MTF50 is approximately 40% of the simulated value. Even small mechanical tolerance might be crucial to MTF. The mechanical assembly will be robust and improved in the final product.

8.2 Analysis of Virtual Reality Headset Measurement Results

The eye relief of the system is 13.4 mm for the left side and 12.1 mm for the right side with interpupillary distance of 64 mm. There is over 1 mm variance between the two. This can be related to the measurement accuracy, since the horizontal centers of the eyebox vary by 2 mm (vertical centers are the same), which can be caused by e.g. a tilt that results in 1 mm offset to opposite directions between the two sides. The exit pupil horizontal and vertical dimensions are 8.4 mm x 13.0 mm and 11.2 mm x 15.9 mm for the left and right side respectively.

The FOV is measured with either the F-theta (i.e. linear barrel distortion) model or the inverse angular mapping function of the HMD Eye. In this thesis, the inverse angular mapping function was used. The results, left and right side respectively, for the horizontal FOV is 79.4° and 79.2° while the vertical FOV is 79.6 and 79.4. The respective results with the F-theta model are slightly higher (a couple of degrees), as predicted from the Figure 8.2. The results of the eyebox and the FOV measurements is collected in Table 8.2.

The analysis results imply that the vertical size of the exit pupil is bigger than the horizontal. If this result is compared to the shape of the lens in Figure 7.7b, the results seem sensible, depending on where exactly is the physical optical axis of the Oculus Go.

The lens' Fresnel circles' center of curvature (i.e. the argued optical axis) appears to be slightly below the spot where the user's nose radix is supposed to be. The horizontal width of the lens is slightly less at the suspected optical axis level in comparison to the nose radix level and also to the vertical width. This might be the cause of the eyebox size difference (that is around 4.5 mm in total), since the lens' width is of the same scale as the eyebox size difference (recall Figure 4.2).

Table 8.2. *Oculus Go eyebox and FOV measurement results. A 95% brightness threshold was used for exit pupil detection.*

Value	**Left side**	**Right side**
Eye relief [mm]	13.4	12.1
Exit pupil horizontal size [mm]	8.4	11.2
Exit pupil vertical size [mm]	13.0	15.9
Interpupillary distance [mm]	64	
Horizontal center difference [mm]	2	
Vertical center difference [mm]	0	
Vertical FOV [degree]	79.6	79.4
Horizontal FOV [degree]	79.4	79.2

The MTF of Oculus Go is shown in Figure 8.5. The MTF was calculated for right and top edges of the square at the proximity of the optical axis (demonstrated in Figure 7.9a). This was done for both left and right eye sides of the DUT.

It is quite intuitive that the MTF of the DUT is worse than that of the HMD Eye. In principle, the MTF of the measurement metrology should not affect the results (or they should taken care of through calibration).

The MTF is almost identical between all four cases apart from the noise manifested at higher frequencies. The original Nyquist and sampling frequencies correspond to the values of 18.4 cycles/deg and 36.8 cycles/deg respectively in this case too. The most important part of the analysis data is within 6 cycles/deg, where the MTF reaches zero.

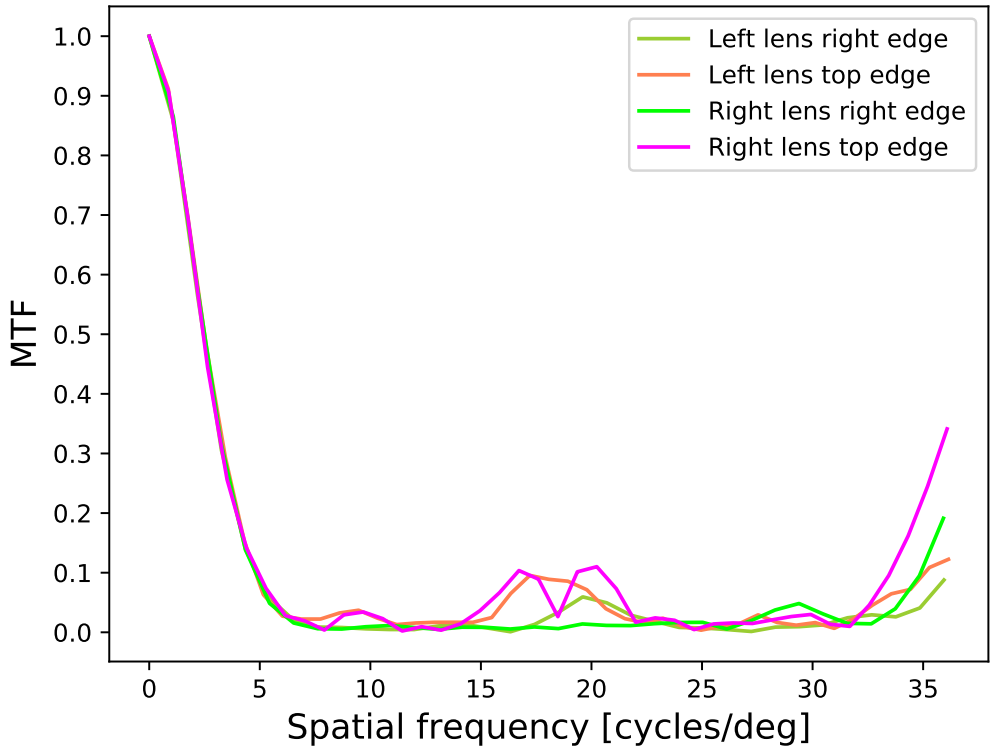


Figure 8.5. *The MTF of the Oculus Go. These results include the MTF of EYE2 and nonidealities, such as shift from the optical axis of the DUT.*

The data within 6 cycles/deg can be considered the real MTF of the system, since higher frequencies include only noise. The MTF50 is approximately 2.5 cycles/deg in all four cases. This is almost one fourth of that of the HMD Eye. Because the MTF of the DUT was measured with white RGB light, chromatic aberration can decrease the MTF. This is on purpose, since the most likely use cases (e.g. pixels producing light of all the three wavelenghts) from the user's point of view are the main focus of the HMD IQ.

The noise peak (magnitude ≈ 0.1) between 16 cycles/deg and 21 cycles/deg is very likely caused by the finite separation of the individual pixels of the DUT. The FOV of the DUT was measured to be approximately 80°, and the sensor size is 1280x1440 (horizontal x vertical pixels). From Figure 7.10, some pincushion distortion is expected from the DUT, i.e. pixels further off the optical axis are stretched transversely. Each pixel thus approximately corresponds to noise frequency of 1280 cycles/80 deg = 16 cycles/deg horizontaly and 1440 cycles/80 deg = 18 cycles/deg. These noise frequency values (which are also analogical to the sampling frequency) fit rather well in the MTF analysis data, at least with the top edges.

The reason why the noise peaks related to finite light emitting area of pixels are not found in the right edge analyses are due to Moire effect, which can be seen in Figure 7.9b. The separation of lit areas are more clear in the case of columns than rows. The effect of dark rows and columns to the MTF excluding the noise has not been investigated yet.

The smallest angular spatial frequency that the DUT can send is half of the frequency that corresponds to the pixel related noise. In this case these frequencies are 16/2 cycles/deg = 8 cycles/deg and 18/2 cycles/deg = 9 cycles/deg. This is the analog of Nyquist frequency and appears e.g. when a given pixel is on, the pixels next to it are off, the pixels next to these are on etc. There appears to be very small signals (magnitude ≈ 0.03) at these frequencies in Figure 8.5. These could be due to small deviations of output power between pixels, but this is hard to tell.

Some noise is apparent at around 35 cycles/deg, just like in the MTFs of the EYE2. In the case of the left side, these noise levels are the same in magnitude, but with the right side they are two times higher at the right edge and three times higher at the top edge. Perhaps this change in the noise can be linked to possible invalid positions of the robot holding the EYE2, but considering the very small differences of the eyebox and FOV properties in Table 8.2, this cannot be done in this case, unfortunately. Keep in mind that there is clearly no substantial differences in the MTFs before 8 cycles/deg, which there should be if the position in the eyebox was severely amiss.

The reason why noise can be seen in the first place in the analysis data is that the source spectrum is not that of an ideal rectangular edge, similar to Equation 3.2. Instead, some frequencies (e.g. those related to finite pixel areas and highest information transmission frequency) are biased, having higher amplitudes than the others. The resolution of the HMD Eye sensor can perhaps also cause tendency to some noise frequencies. The DFT of this kind of signal surely is not that of an ideal MTF. The noise levels were lowered with the Hamming window in the MTF algorithm, but they can still appear.

The relative geometric distortion describes how much the image points are dislocated in comparison to the points close to the optical axis. This is calculated as follows. First, the dots are detected in the digital image and then the dot closest to the center of the image is fetched. After this, the four closest dots to the center dot (top, right, bottom and left) are used to calculate the undistorted locations (i.e. location without distortion) of all the dots, since these four dots are in practise not distorted (remember the dot grid in Figure 7.10 consists of 391 dots). The angles are calculated from pixel locations using the F-theta lens model. Finally the relative angular distortion is calculated with Equation 8.1:

$$K_{\text{dist}} = \frac{\theta_{\text{dist}} - \theta_{\text{undist}}}{\theta_{\text{undist}}}, \tag{8.1}$$

where θ_{dist} and θ_{dist} are the polar angles of dot objects (with and without the distortion of the DUT) respective to the HMD Eye's object space. This analysis of Figure 7.10 with a third degree polynomial fit is shown in Figure 8.6.

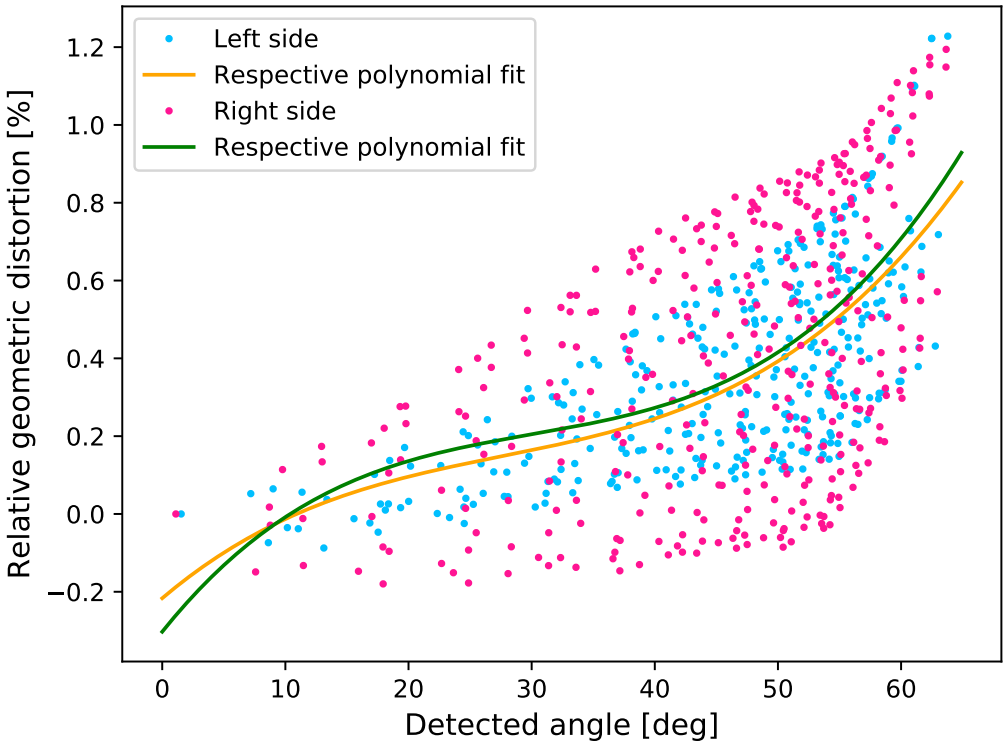


Figure 8.6. *Oculus Go's relative distortion of dots.*

The results in Figure 8.6 show that the distorted polar angles are higher than the relative undistorted polar angles. For example, at the polar field of 45° (where the screen can still be seen in all azimuth angles) undistorted angle would be ≈ 44.8°. At higher polar fields, i.e. from the corners of the screen, the distortion rapidly increases. Ergo, the Oculus Go definitely has pincushion distortion. The fitted model does not represent image points close to the optical axis that well.

Chromatic aberration considered in the analysis is the relative lateral chromatic aberration of the image points, i.e. how much the angular coordinates of the image points at the NED's virtual image plane vary in percentages in respect to some comparison wavelength, demonstrated in Equation

$$K_{\text{ca}} = \frac{\theta_{\text{meas}} - \theta_{\text{ref}}}{\theta_{\text{ref}}}, \tag{8.2}$$

where θ_{meas} is the angular coordinate of the measured wavelength while θ_{ref} is the angular coordinate of the wavelength it is compared to. The comparison wavelength is green, since it is in the middle of the wavelength spectrum of light in comparison to blue and red. The analysis results are shown in Figure 8.7. The measurement data is similar to that of Figure 7.10, but measured separately with each of the three wavelengths.

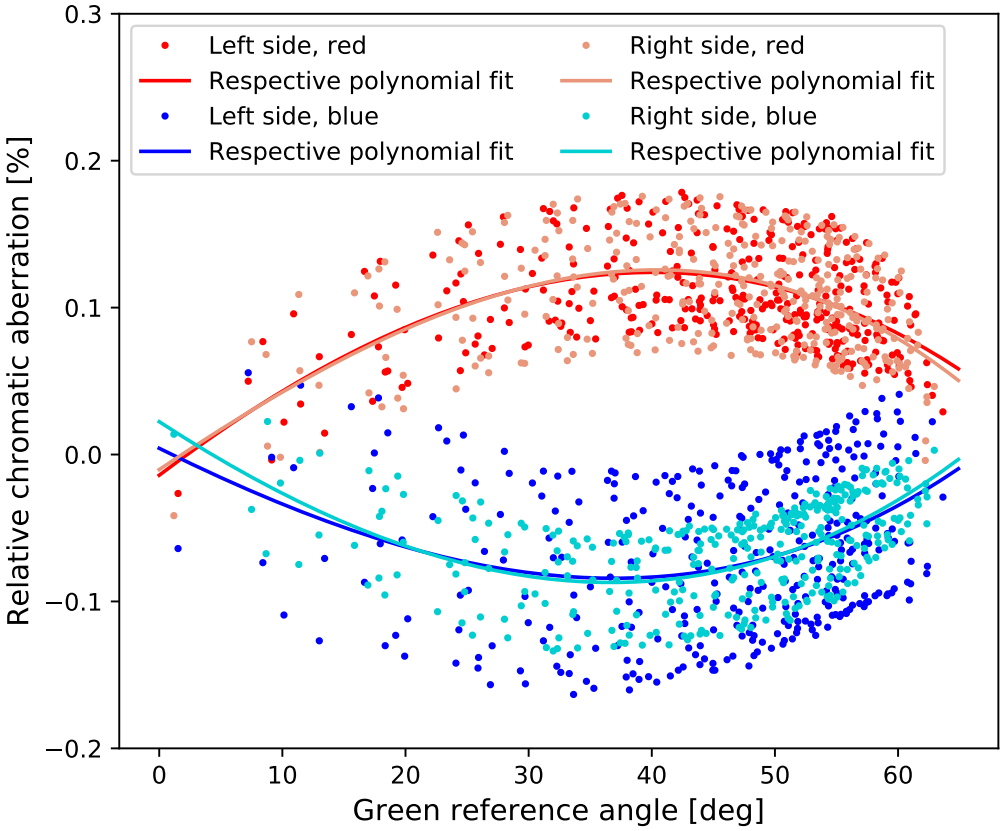


Figure 8.7. *Oculus Go's relative transverse chromatic aberration of dots.*

The results show that red image points are pincushion distorted the most. However, the image points seem to be located at the same angluar coordinate at the limit of the DUT FOV. Closer look at the measurement data does not frankly endorse this at every corner of the screen (in some corners the separation of the wavelength specific image points seems to increase). There might be some difficulty detecting the center of the dots, caused by e.g. the chromatic aberration (the wavelength specific FWHM is most likely quite large).

The fitted models for both sides of the DUT are much alike with the highest error being 0.025% at small angles with blue light. There is also much less variance between the data points and the fitted model in the case of the right side with blue light. This at least implies there is other difficulties within the analysis than just the chromatic aberration of one wavelength (or rather a band centered around this wavelength), since the light sources are similar on both sides of the DUT.

9 CONCLUSION

The results of this thesis can be divided into three categories: research and pondering of the physical principles causing the necessity of calibration and characterization, the developed and enhanced methods enabling this and some example results for one HMD Eye and for a virtual reality headset. This thesis is appropriate additional material for the HMD Eye calibration standard operation procedure (SOP).

The theory section has provided new tools in the HMD Eye project of which some are already implemented in the project. The advantages and suitability of the MTF analysis algorithm, provided by ISO-12233 standard, have been analysed. The effects caused by noise and aliasing are diminished by supersampling into given number of bins with a slanted edge square target of given rotation. This describes the real properties of the optics. Apprehension of this subject enabled the more sophisticated design of the focus calibration and its automation design. The problems associated with a single focus and calibration of this kind of system are laid out and supported with the measurement results.

Geometrical distortion has been modeled more extensively than it had been at the start of the project. As a result, the angular mapping function of the imaging system can be derived on the basis of distortion calibration alone alongside the definition of the optical axis. In other words, no sphere-like targets for separate angular calibration are deemed necessary, depending on the allowed error margin. The fact that not all measurements benefit from distortion correction is remarked. Tangential distortion has given insight about the quality of the system. This indicates that there is consensus between the imaging performance and the nonideal forms of distortion. This argument is based on distortion calibration of three different HMD Eyes of which two have had somewhat worse imaging performance.

The nature of the eyebox and optical properties of near eye displays and head mounted displays has been studied based on theory of optics and real system properties (in-house measurements and patents by companies producing VR and AR devices). Due to its small entrance pupil among other features, HMD Eye is a very ideal solution for analysis of these properties.

By using the test unit (’zero batch’) HMD Eye, several discoveries were found. These results support and endorse the contents of this thesis. The analysis data format for the HMD Eye has now been documented. This will help with software usage and HMD Eye specifications.

It is hard to tell which direction the head mounted display industry is going. Consumer interest is becoming more visible the more options are available for public markets, of which AR devices currently are really not.

HMDs are yet to go through enormous leaps in term of immersion: vergence-accommodation conflict has to be solved and field of view has to match the human eye. HMD Eye needs to be ready for these evolutions, which it is since motorized focus is on its way and its FOV is already high to begin with. Competitors might also become more inclined towards very high FOVs and motorized sensors.

REFERENCES

[1] *AR/VR Headsets Return to Growth in the First Quarter As New Models and Use Cases Restore Momentum to the Market, According to IDC*. URL: https://www.idc.com/getdoc.jsp?containerId=prUS45326719 (visited on 11/19/2019).

[2] W. Steen. *Principles of Optics M. Born and E. Wolf, 7th (expanded) edition, Cambridge University Press, Cambridge, 1999, 952pp. 37.50/US $59.95, ISBN 0-521-64222-1*. 2000. DOI: 10.1016/S0030-3992(00)00061-X.

[3] D. W. Ball. *Field Guide to Spectroscopy*. 2009. DOI: 10.1117/3.682726.

[4] E. Hecht. *Optics 4th edition*. 1998. ISBN: 9780321188786. DOI: 10.1119/1.3274347.

[5] Field guide to geometrical optics. *Choice Reviews Online* (2013). ISSN: 0009-4978. DOI: 10.5860/choice.42-1002.

[6] J. Melzer, F. T Brozoski, T. Letowski, T. Harding and C. E Rash. Guidelines for HMD design. *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues* (May 2009).

[7] Greg Hollows. *Depth of Field and Depth of Focus | Edmund Optics*. URL: https://www.edmundoptics.com/resources/application-notes/imaging/depth-of-field-and-depth-of-focus/ (visited on 06/07/2019).

[8] International Organisation for Standardization. Photography - Electronic still picture imaging — resolution and spatial frequency responses. *ISO 12233:2014* (2014).

[9] N. Storey. *Electronics. A Systems Approach*. 4th edition. Pearson Education Canada, 2009.

[10] G. D. (D. Boreman. *Modulation Transfer Function in Optical and Electro-Optical Systems*. 1st. SPIE Press, 2001, 110. ISBN: 9780819441430.

[11] *Digital Signal Processing: Fundamental and Applications*. Elsevier, 2008.

[12] *Imaging Electronics 101: Camera Resolution for Improved Imaging System Performance*. URL: https://www.edmundoptics.com/resources/application-notes/imaging/camera-resolution-for-improved-imaging-system-performance/ (visited on 06/10/2019).

[13] J. Weng, P. Coher and M. Herniou. Camera Calibration with Distortion Models and Accuracy Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1992). ISSN: 01628828. DOI: 10.1109/34.159901.

[14] J. Wang, F. Shi, J. Zhang and Y. Liu. A new calibration model of camera lens distortion. *Pattern Recognition* (2008). ISSN: 00313203. DOI: 10.1016/j.patcog.2007.06.012.

[15] *OpenCV: Camera Calibration and 3D Reconstruction*. URL: https://docs.opencv.org/4.1.0/d9/d0c/group%7B%5C_%7D%7B%5C_%7Dcalib3d.html (visited on 07/04/2019).

[16] D. A. Kerr. Derivation of the “Cosine Fourth” Law for Falloff of Illuminance Across a Camera Image Douglas. *dougkerr.net/Pumpkin/articles/Cosine_Fourth_Falloff.pdf* (2007).

[17] M. Aggarwal, H. Hua and N. Ahuja. On cosine-fourth and vignetting effects in real lenses. 2002. DOI: 10.1109/iccv.2001.937554.

[18] *ON Semiconductor PYTHON 16K CMOS sensor*. URL: https://www.onsemi.com/pub/Collateral/NOIP1SN025KA-D.PDF (visited on 09/30/2019).

[19] Canada-France-Hawaii Telescope (CFHT). *Flat field and dark frame correction*. URL: http://www.cfht.hawaii.edu/%7B~%7Dbaril/Pyxis/Help/flatdarkfield.html (visited on 06/28/2019).

[20] C. Greb, P. Laskey and K. Schwap. *Introduction to Digital Camera Technology*. URL: https://www.leica-microsystems.com/science-lab/introduction-to-digital-camera-technology/ (visited on 10/24/2019).

[21] R. Hui and M. O'Sullivan. *Fiber Optic Measurement Techniques*. 2009. ISBN: 9780123738653. DOI: 10.1016/B978-0-12-373865-3.X0001-8.

[22] E. M. V. Association. EMVA Standard 1288, Standard for Characterization of Image Sensors and Cameras. *Release* (2010).

[23] R. Navarro. *The optical design of the human eye: A critical review*. 2009. DOI: 10.3921/joptom.2009.3.

[24] G. Kramida. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics* (2016). ISSN: 10772626. DOI: 10.1109/TVCG.2015.2473855.

[25] G. K. Aguirre. The entrance pupil of the human eye. *bioRxiv* (2018). DOI: 10.1101/325548.

[26] *Wikipedia: Comparison of Virtual Reality Headsets*. URL: https://en.wikipedia.org/wiki/Comparison_of_virtual_reality_headsets (visited on 06/10/2019).

[27] B. C. Kress and W. J. Cummings. Towards the ultimate mixed reality experience: Hololens display architecture choices. *Digest of Technical Papers - SID International Symposium*. 2017. DOI: 10.1002/sdtp.11586.

[28] O. Cakmakci and J. Rolland. *Head-worn displays: A review*. 2006. DOI: 10.1109/JDT.2006.879846.

[29] J. D. Perreault, X. Hu, Y. Qin and S. Bierhuizen. *Optical Field Curvature Control using Multi-layer Fresnel Lens in VR Display*. U.S. Patent 10215890 B2, Jul. 2019. URL: https://pdfpiw.uspto.gov/.piw?Docid=10215890.

[30] E. M. Richards. *Angle Selective Filter for Near Eye Displays*. U.S. Patent 0212482 A1, Jul. 2019. URL: http://pdfaiw.uspto.gov/.aiw?Docid=20190212482.

[31] *PlayNitride microLED*. URL: https://www.youtube.com/watch?v=qH5AFlTM0fQ (visited on 07/15/2019).

[32] J. Y. Lee and S. I. Yoo. Automatic detection of region-mura defect in TFT-LCD. *IEICE Transactions on Information and Systems* E87-D.10 (2004), 2371–2378. ISSN: 09168532.

[33] P. K. Uttwani, B. C. Villari, K. N. Narayanan Unni, R. Singh, A. Awasthi and Deepak. Detection of physical defects in full color passive-matrix OLED display by image driving techniques. *IEEE/OSA Journal of Display Technology* 8.3 (2012), 154–161. ISSN: 1551319X. DOI: 10.1109/JDT.2011.2168805.

[34] Z. Wang, J. Gao, C. Jian, Y. Cen and X. Chen. OLED Defect Inspection System Development through Independent Component Analysis. *TELKOMNIKA Indonesian Journal of Electrical Engineering* 10.8 (2013), 2309–2319. ISSN: 2302-4046. DOI: 10.11591/telkomnika.v10i8.1968.

[35] I. O. for Standardization. Optics and Optical instruments - Veiling glare of image-forming Systems - Definitions and methods of measurement. *ISO 9358:1994* (2014).

[36] *Allied Vision Prosilica GT 4096 Data sheet*. URL: https://www.alliedvision.com/en/products/cameras/detail/Prosilica%5C%20GT/4096/action/pdf.html (visited on 09/30/2019).

[37] *Allied Vision Manta G-1236B Data sheet*. URL: https://www.alliedvision.com/en/products/cameras/detail/Manta/G-1236/action/pdf.html (visited on 09/30/2019).

[38] *Sony IMX304 Data sheet*. URL: https://www.1stvision.com/cameras/sensor_specs/IMX304.pdf (visited on 09/30/2019).

[39] A. B. Watson. A formula for the mean human optical modulation transfer function as a function of pupil size. *Journal of Vision* (2013). ISSN: 1534-7362. DOI: 10.1167/13.6.18.

[40] *TRIOPTICS ImageMaster*. URL: https://www.trioptics.com/products/mtf/imagemaster-vr-ar/ (visited on 07/26/2019).

[41] *Gamma Scientific NED-LMD Near-Eye Display Measurement System*. URL: https://www.gamma-sci.com/products/ned-lmd-near-eye-display-measurement-system (visited on 07/26/2019).

[42] *Radiant Vision Systems AR/VR Lens*. URL: https://www.radiantvisionsystems.com/products/arvr-lens?utm_campaign=arvr-lens&utm_medium=pr&utm_source=pr (visited on 07/26/2019).

[43] British Standard Institute. BS ISO 14524:2009 Photography - Electronic still-picture cameras - Methods for measuring opto-electronic conversion functions (OECFs). *BSI Standards Publication* (2009).

[44] *Leading Optical Product Design Software for Engineering Teams - Zemax*. URL: https://www.zemax.com/ (visited on 06/11/2019).

[45] *Image Engineering CAL2*. URL: https://www.image-engineering.de/products/equipment/illumination-devices/638-cal2 (visited on 06/28/2019).

[46] *Integrating spheres, RP-photonics*. URL: https://www.rp-photonics.com/integrating_spheres.html (visited on 07/17/2019).

[47] J. Hovila. *NEW MEASUREMENT STANDARDS AND METHODS FOR PHOTOMETRY AND RADIOMETRY Doctoral Dissertation*. 2005. ISBN: 9512279541.

[48] OMG. *OMG Unified Modeling Language (OMG UML), Version 2.5*. 2010. DOI: 10.1007/s002870050092.

[49] S. Brown. *The C4 model for visualising software architecture*. URL: https://c4model.com/ (visited on 06/07/2019).

[50] J. Schanda. *Colorimetry: Understanding the CIE System*. 2007. ISBN: 9780470049044. DOI: 10.1002/9780470175637.

[51] *Oculus Go website*. URL: https://www.oculus.com/go/?locale=fi_FI (visited on 10/15/2019).

[52] *AliExpress AR Electrics Store*. URL: https://www.aliexpress.com/item/33009974801.html (visited on 10/24/2019).

[53] M. Sonka, V. Hlavac and R. Boyle. *Image Processing, Analysis, and Machine Vision*. 2008. ISBN: 049508252X.

[54] J. Ijspeert, T. V. D. Berg and H. Spekreijse. An improved mathematical description of the foveal visual point spread function with parameters for age, pupil size and pigmentation. *Vision Research* 33.1 (1993), 15–20. ISSN: 0042-6989. DOI: 10.1016/0042-6989(93)90053-Y.

[55] L. N. Thibos, X. Hong, A. Bradley and X. Cheng. Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *J. Opt. Soc. Am. A* 19.12 (Dec. 2002), 2329–2348. DOI: 10.1364/JOSAA.19.002329.